Economic Modelling 91 (2020) 455–462

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.journals.elsevier.com/economic-modelling

Cross market predictions for commodity prices

Shusheng Ding, Yongmin Zhang *

School of Business and Research Academy of Belt and Road, Ningbo University, China

ARTICLE INFO

JEL classification:
E30
F00

Keywords:
Commodity futures
Commodity market liquidity
Cross-market cointegration
Lon-run price equilibrium

ABSTRACT

This paper shows that commodity prices can be predicted from cross-market information by establishing long-run cross-market commodity price equilibrium models, which are characterized by a linear relation between prices across different markets. Using data from five representative commodity markets (oil, copper, gold, corn, and cattle) during the period 2005–2018, we demonstrate that oil and industrial metal markets have formed a long-run price equilibrium with other markets across different commodity families. However, agriculture and gold markets do not tend to have long-run price equilibrium relations with other commodity markets. Furthermore, we show that the absence of a price equilibrium is due to the cross-market liquidity interference effect. After we control for the liquidity effect, long-run cross-market commodity price equilibrium relations are reestablished for agriculture and gold markets. These results can aid in demonstrating that liquidity can capture most of the missing information that is not reflected in price dynamics in less liquid markets, such as agriculture and gold markets. Therefore, less liquid commodity price predictions require both prices and liquidity levels from cross-markets, while liquid commodity prices (oil and metal) can be predicted based solely on cross-market prices.

1. Introduction

There is consensus in the literature that commodity price movements preserve a strong relation with macroeconomic variables, which are often used to predict commodity price moves (Garner, 1989; Lof and Nyberg, 2017). However, such predictive models are usually only valid for actively traded commodities, such as energy and metal, which are sensitive to macroeconomic conditions. Therefore, price predictions for less actively traded commodity futures, such as agricultural futures, remain unresolved. If we can establish cross-market price equilibrium relations, then less liquid commodity price moves could be predicted from more liquid commodity futures, the price dynamics of which can be tracked instantaneously from actively traded futures markets. Nevertheless, intensive research on cross-market price equilibrium relations is scarce, in which the cross-market price equilibrium can be defined to mean that the commodity price is linked to other cross-market commodity prices through a linear relation. Accordingly, in this paper, we address the following research questions:

(1) Which commodities have formed long-run cross-market price equilibrium relationships with other commodity markets?

(2) For those commodities for which prices do not tend to establish a cross-market equilibrium, what are the interfering factors, and how might their prices be predicted?

Cross-market equilibrium relations could attract great attention from both governments and practitioners, because the price moves of one commodity can be predicted by price moves from other commodity markets. Because of the price interactions between different commodity markets, a cross-market price equilibrium can assist policymakers in monitoring inflation dynamics more effectively and thus formulate adaptive policies in a timely manner, thereby protecting the economy contemporaneously. To forecast the long-run inflation level, we need aggregate price levels across different commodity markets. Thus, we must predict prices not only for actively traded (liquid) commodities but also for less actively traded (illiquid) commodities. If liquid and less liquid commodities' long-run prices satisfy cross-market equilibrium relations, we could predict long-run prices for less liquid commodities from the long-run prices of liquid commodities, which are traded in efficient markets and are relatively easy to forecast based on efficient market hypothesis theory (Chordia et al., 2008; Iwatsubo et al., 2018). Therefore, a cross-market equilibrium for the long-run price level could effectively predict the long-run inflation level. Such a cross-market price equilibrium can also facilitate commodity futures traders in designing

* Corresponding author.
E-mail address: zhangyongmin@nbu.edu.cn (Y. Zhang).

https://doi.org/10.1016/j.econmod.2020.06.019
Received 2 July 2019; Received in revised form 27 June 2020; Accepted 27 June 2020
Available online 7 July 2020

more useful cross-market arbitrage and hedge strategies. In the case of arbitrage, if the current market price for a commodity is below the long-run equilibrium price predicted by the cross-market equilibrium model, we can take a long position on this commodity's futures and sell it when the price reaches the equilibrium price in the future and, in turn, realize a risk-free profit. For hedging applications, if one commodity market is volatile, we can take futures positions in other commodity markets to hedge against turmoil.

The starting point described above notably distinguishes our work from the literature, and our contributions are threefold. First, we use cross-market information to predict commodity price moves, while previous studies mainly use macroeconomic variables to predict commodity prices. Furthermore, because most macroeconomic variables are only available on a low-frequency basis (such as monthly or quarterly), prior studies can only employ low-frequency data. In contrast, our data cover both price and liquidity information, which are available at a higher—-even daily—frequency. A number of prior studies emphasize the importance of employing high-frequency data in policy effect analysis (Coleman, 1990; Serletis, 1994; Westgaard et al., 2011; Zhang et al., 2019). Finally, our study's most significant contribution is identifying the role of liquidity in long-run price equilibria across different commodity markets, which offers new insights into the literature on commodity price prediction.

The integral relation between commodity prices and economic variables has been well studied. For instance, Cody et al. (1991) delineate the close relationship between commodity prices and the consumer price index (CPI) as a key economic indicator. Chinn and Coibion (2014) show that understanding movements and changes in commodity prices can also be helpful in near-future policy formulation. More recently, numerous studies have demonstrated that commodity prices play a vital role in affecting macroeconomic indicators such as inflation and economic output. Mallick and Sousa (2013) estimate a battery of modern dynamic macroeconometric models to underscore the importance of commodity price shocks, which lead to a rise in inflation and require more aggressive behavior from central banks with respect to inflation stabilization. Similarly, Holtemöller and Mallick (2016) employ both an estimated VARX and a small open-economy New Keynesian model to demonstrate that global food price shocks constitute an important aspect of inflation in India. Alquist et al. (2019) find that global economic fluctuations have been substantially influenced by commodity price shocks.

The cointegration method has been widely used to analyze the long-run equilibrium relationship between commodity prices and economic activities. For instance, Lardic and Mignon (2008) reveal the asymmetric cointegration relation between oil prices and economic activities. However, prior studies examine the price cointegration relation between economic variables and commodity prices to show a long-run equilibrium relation for predicting price movement trends (Berenguer-Rico and Gonzalo, 2014). To explore *long-run cross-market price equilibrium relations,* we study commodity price *cointegration relationships with both individual commodity markets and the aggregate global commodity index on a higher frequency (daily) basis. The cointegration method has proven to be a powerful tool in forecasting variables in one market based on other markets (*Hassan, 2003*). Furthermore, multiple studies show the existence of liquidity commonalities in both stock and commodity markets (*Chordia et al., 2000; Korajczyk and Sadka, 2008; Marshall et al., 2013; Frino et al., 2014*). Therefore, we also incorporate liquidity variables in our cointegration study of commodity prices, which complement commodity prices in* our cointegration analysis.

To investigate commodity price relations, we conduct cointegration tests across different commodity futures markets, represented by cattle and corn in agricultural markets, copper in the metal market, oil in the energy market, and the gold market. Cointegration is a stronger relation between prices in different markets than either correlation or Granger causality. Cointegration yields a structured linear equation that can predict a commodity's price from other commodity prices. Since the prices of commodity futures tend to comove and the liquidities of different markets exhibit commonality (Zhang et al., 2019*)*, we could argue that commodity prices tend to be cointegrated. However, the results from our cointegration tests suggest non-cointegrated relations even after confirming Granger causal relations between agriculture and gold markets. Another surprising piece of evidence we report is that when we control for liquidity variables in regression equations, the residuals of cointegration become stationary. These results are complementary to arguments that liquidity contains most of the market noise from the perspective of commodity futures markets (Koski and Michaely, 2000; Sockin and Xiong, 2015). Individual commodity prices can be predicted from other markets only if we include liquidity measures as additional prediction variables. More importantly, our study demonstrates that the liquidity level can be considered an interference factor when the *long-run price equilibrium* relationship between commodity prices evaporates.

In our study of cointegration with the aggregate commodity market, which is represented by the global commodity CRB price index, we find that the oil and copper markets are cointegrated with the global market, while agriculture and gold prices are not driven by the global CRB index. This finding might be explained by different market characteristics for energy and metal, which are strongly linked to the global economy. On the other hand, agricultural markets are strongly influenced by government regulation, and the gold market is driven by investment and hedging activities.

The paper is organized as follows. In Section 2, we describe the data and a suitable liquidity measure for commodity markets. In Sections 3 and 4, respectively, we present our empirical results and discuss the liquidity effects on futures price cointegrations with both cross-sectional markets and the aggregate global commodity market. Section 5 concludes.

2. Data and methodology

According to the NASDAQ commodity family report (2013)*, there are five families of commodities, namely, energy (such as crude oil, (oil hereafter)), agricultural (such as corn), livestock (such as live cattle), precious metals (such as gold) and industrial metals (such as copper). We select one commodity from each family (as indicated in the brackets) to construct a cross-sectional commodity portfolio. The representative commodities are the most actively traded commodities. The trading volumes of crude oil, corn, copper, live cattle and gold are the highest in each section (*Kowalski, 2014*). The selected commodity prices are futures prices with the nearest maturities of the particular commodity.*

For the data description, the superscript 'ca' represents live cattle commodity futures, 'cop' represents copper commodity futures, 'cor' represents corn commodity futures, 'g' represents gold commodity futures, 'o' represents oil commodity futures, and 'ML' represents market liquidity indicator. 'L' stands for the normalized roll measure of liquidity, 'P' stands for the commodity price and 'r' stands for the realized returns of commodity futures. The data provider is Thomson Datastream, and the data period is from Jan 1, 2005 to 31 Dec 2018, which is the common period for maximal available data; in addition, all data are collected on a daily basis.

We first define the commodity liquidity measure (L_t) adopting the bid-ask spread. In our paper, we use the effective spread estimator developed by Roll (1984), which has been utilized in a number of financial studies such as Goyenko et al. (2009) and Corwin and Schultz (2012). The proxy utilizes the autocovariance of the daily price changes as an effective measure of the bid-ask spread. Roll's starting point is that the traded assets have fundamental value, denoted by V_t, as:

$$V_t = V_{t-1} + \eta_t, \tag{1}$$

where η_t reflects new information arrival, which is assumed to be independent of the previous period information under the efficient market hypothesis.

Next, Roll (1984) denotes S_t as the last observed trade price on day t and assumes that S_t follows the following process:

$$S_t = V_t + \frac{1}{2} EQ_t, \tag{2}$$

where E is the effective spread and Q_t is a buy/sell indicator for the last trade that equals +1 for a buy and −1 for a sell. He further assumes that Q_t is equally likely to be +1 or −1 and Q_t is also serially uncorrelated and independent of η_t.

Then, he takes the first difference of (2) and plugs in the result from equation (1), which yields:

$$\Delta S_t = \frac{1}{2} E \Delta Q_t + \eta_t,$$

where ΔQ_t measures the change of the order type from two consecutive days, and Δ is the change operator, namely, $\Delta Q_t = Q_t - Q_{t-1}$ (Goyenko et al., 2009).

As a result, $Cov(\Delta S_t, \Delta S_{t-1}) = -\frac{1}{4}E^2$, or equivalently, $E = spread = 2\sqrt{-\mathrm{cov}(\Delta S_t, \Delta S_{t-1})}$.

However, as the autocovariance is positive, the formula is undefined. Therefore, we use a modified version of the Roll estimator (Goyenko et al., 2009):

$$l_t = spread = \begin{cases} 2\sqrt{-\mathrm{cov}(\Delta S_t, \Delta S_{t-1})}, & \mathrm{cov}(\Delta S_t, \Delta S_{t-1}) \le 0 \\ 0, & \mathrm{cov}(\Delta S_t, \Delta S_{t-1}) > 0 \end{cases}. \tag{3}$$

For empirical liquidity measures, the statistical summary for liquidity measures in five representative commodity futures markets is presented in Table 1, and the liquidity measures summary for the other eight commodity futures markets is presented in Table 1, which are cocoa (l^{coa}_t), coffee (l^{cof}_t), natural gas (l^{ng}_t), orange juice (l^{oj}_t), silver (l^{s}_t), soybeans (l^{so}_t), sugar (l^{su}_t) and wheat (l^{w}_t) commodity markets, respectively.

This table gives a detailed statistical summary of the remaining eight liquidity measures, and it can be seen that there would also be size effects across different commodity futures markets.

From Table 1, we can see that corn and gold futures liquidities have the first and second largest mean values, implying that these two futures are less liquid. For cattle futures, the low mean value results from the large proportion of zero volume trading days, which is also an indicator of illiquidity (see Liu, 2006). According to tbl1Tables 1 and 2tbl2, it may also be arguable that there could be size bias effects among different futures markets since the means and standard deviations of liquidity measures are widely divergent across different commodity futures markets. To eliminate the size bias effects, we normalize our liquidity measures, and the normalized liquidity measure is defined as:

L_t= (liquidity measure-mean of the liquidity measure)/standard deviation of the liquidity measure.

This table gives a detailed statistical summary for five representative normalized liquidity measures, and it can be seen that the size effect across different commodity futures markets has been effectively eliminated.

Table 3 and Table 4 show the normalized liquidity measures data summary for all 13 commodity futures markets, and it is convincing that size bias effects are effectively eliminated. All regression results in this paper are based on normalized liquidity measures. Moreover, since the daily liquidity measure is extremely noisy, to smooth the time series process of liquidity measures, we adopt a rolling over strategy to calculate the liquidity measures based on the normalized data. In other words, we obtain the moving-average daily Roll measure, which indicates that one particular day's liquidity measure is effectively the average of the previous month's (the past 21 days) liquidity measures, and thereby our liquidity data frequency is on a daily basis. Let $\overline{L}_t$ be the rolling average of liquidities from the past 21 trading days, i.e.,

$$\overline{L}_t = \frac{1}{21} \sum_{i=0}^{20} L_{t-i}$$

where L_{t-i} is the liquidity measure at day t-i.

In addition, we adopt the CRB Index as the proxy for the commodity futures market price index and returns. The CRB Index is currently composed of 19 commodities, including energy, agriculture and metal, which could be a good representation of overall commodity prices. However, only 13 of 19 commodity futures data are available. Therefore, in addition to the individual liquidity measure, we define the market liquidity measure (ML_t) as the equally weighted average of 13 commodity futures liquidity measures, which are live cattle (L^{ca}_t), cocoa (L^{coa}_t), coffee (L^{cof}_t), copper (L^{cop}_t), corn (L^{cor}_t), crude oil (L^{o}_t), gold (L^{g}_t), natural gas (L^{ng}_t), orange juice (L^{oj}_t), silver (L^{s}_t), soybeans (L^{so}_t), sugar (L^{su}_t) and wheat (L^{w}_t). Namely, the market liquidity (ML_t) can be denoted

Table 1
Statistical summary table for liquidity measures in five representative commodity futures markets. This table gives a detailed statistical summary of the original five representative liquidity measures, and it can be seen that there might be a large size effect across different commodity futures markets.

	Obs	Mean	Std. Dev.	Min	Max
l^{ca}_t	2852	0.38	0.92	0	7.01
l^{cor}_t	2852	3.74	5.50	0	32.44
l^{cop}_t	2852	0.03	0.03	0	0.21
l^{g}_t	2852	5.47	6.91	0	37.95
l^{o}_t	2852	1.53	1.91	0	11.94

Table 2
Statistical summary table for liquidity in the other eight commodity futures markets.

	Obs	Mean	Std. Dev.	Min	Max
l^{coa}_t	2852	38.27	108.57	0	1089.54
l^{cof}_t	2852	2.53	6.06	0	60.64
l^{ng}_t	2852	0.072	0.093	0	0.48
l^{oj}_t	2852	2.24	8.24	0	85.22
l^{s}_t	2852	0.19	0.25	0	1.72
l^{so}_t	2852	7.56	9.29	0	45.45
l^{su}_t	2852	0.19	0.25	0	1.75
l^{w}_t	2852	5.78	7.20	0	44.86

Table 3
Statistical summary table for normalized liquidity measures.

	Obs	Mean	Std.Dev.	Min	Max
L^{ca}_t	2852	0.0001	1.00	−0.41	7.18
L^{cor}_t	2852	0.0001	1.00	−0.68	5.21
L^{cop}_t	2852	0.0001	1.00	−0.92	5.53
L^{g}_t	2852	0.0001	1.00	−0.79	4.69
L^{o}_t	2852	0.0001	1.00	−0.80	5.44

Table 4
Statistical summary table for normalized liquidity measures in the other eight commodity futures markets. This table gives a detailed statistical summary for the remaining eight normalized liquidity measures, and it can be seen that the size effect across different commodity futures markets has been effectively eliminated.

	Obs	Mean	Std. Dev.	Min	Max
L^{coa}_t	2852	−0.002	1.00	−0.35	9.68
L^{cof}_t	2852	−0.001	1.00	−0.41	9.58
L^{ng}_t	2852	−0.0001	1.00	−0.77	4.43
L^{oj}_t	2852	−0.0001	1.00	−0.27	10.0
L^{s}_t	2852	−0.0001	1.00	−0.74	5.92
L^{so}_t	2852	0.00001	1.00	−0.81	4.07
L^{su}_t	2852	−0.0001	1.00	−0.77	6.06
L^{w}_t	2852	−0.001	1.00	−0.80	5.42

as:

$$ML_t = \frac{1}{13}\sum_{i=1}^{13}\left(\overline{L}_t^i\right)$$

where i = 1, 2, …, 13, representing live cattle, cocoa, coffee, copper, corn, crude oil, gold, natural gas, orange juice, silver, soybeans, sugar and wheat, respectively.

For the empirical analysis, we use commodity prices (P^i_t) from commodity futures markets, and we normalize the commodity prices by taking the natural log of all commodity prices. Additionally, we also use the realized returns of different commodity futures markets, denoted as r^i_t, which is defined as $r_t^i = \ln\left(\frac{P_t^i}{P_{t-1}^i}\right)$. As a result, the realized returns, the log commodity prices and the monthly moving-average of normalized liquidity measures are three key variables for the empirical analysis in this paper.

Furthermore, to ensure that the cointegration test is an appropriate method for the analysis of this paper, we conduct the Dicky-Fuller test for the individual price series. The test results are presented in Table 5, where most of the price series are nonstationary. The cointegration test would be particularly helpful in tackling the analysis of the relationship between nonstationary time series (Serletis, 1994). Therefore, the cointegration test would be a sound method for dealing with those nonstationary time series for further analysis of the residual part.

3. Individual liquidity effect on cross-sectional prices cointegration

Commodities have firm correlations with each other, and it is therefore arguable that the five commodity prices tend to be cointegrated (see Fig. 1). Evidence could also be found in the literature; for instance, Zhang et al. (2010) find a cointegration relationship between gold and oil futures markets. As a consequence, it is reasonable to claim that there is a cointegration relationship among different commodity futures markets. Therefore, we implement a series of cointegration tests to test the cointegration relationship.

The cointegration test was developed by Granger (1986) and Engle et al. (1987). The series must be differenced d times to achieve stationarity; then, the series is known as I(d). According to Engle et al. (1987), the residuals of the regression of the two nonstationary series are stationary, and there may be a cointegration relationship between the two series. It is also true for multiple nonstationary series (Engsted et al., 1997). As a result, we use the regression model to obtain the residuals for each market, and then we use the Dicky-Fuller test to determine whether the residuals of the five futures markets are stationary. If they are stationary, we can argue that the five different commodity futures markets are cointegrated. The regression results of the cointegration test regarding five different commodity futures markets are presented in Table 6. The DF statistics for the Dicky-Fuller stationarity test for residuals for the system of regression equation (4) are presented in Table 7. We found that only the oil market and copper market are strongly cointegrated with other markets. The rest of the commodity markets do not seem to be cointegrated with other markets or not strongly enough, only at the 10% significance level.

Table 5
Dicky-Fuller test results for the individual price series. Significance levels are denoted by *, **, ***, which correspond to the 10%, 5%, and 1% levels, respectively.

	$\ln P^{ca}_t$	$\ln P^{cor}_t$	$\ln P^{cop}_t$	$\ln P^{g}_t$	$\ln P^{o}_t$
DF	−1.51	−1.38	−1.95	−1.38	−3.12**
	(0.69)	(0.58)	(0.30)	(0.59)	(0.03)

$$\ln P_t^i = \alpha_i + \sum_{j=1, j\neq i}^{5} \beta_i \ln P_t^j + \varepsilon_{i,t} \tag{4}$$

where i = 1, 2, 3, 4, 5, representing ca, cor, cop, g, o, respectively.

This table reports the Dicky-Fuller test for residuals from the regression model equation (4). From the p-values indicated in the brackets, some of the residuals are not stationary, such as cattle and gold. Significance levels are denoted by *, **, ***, which correspond to the 10%, 5%, and 1% levels, respectively, and p-values are reported in brackets.

The results may stem from the issue presented in Granger (1986). Granger (1986) argued that the prerequisite of cointegration is that there must be Granger causality in at least one direction for the two series, as one variable can help forecast the other. As a result, we conduct a series of Granger tests to see whether there is a Granger causality for the three commodity prices (i.e. cattle, corn and gold) that display the results of the cointegration tests that are insignificant or that are not strongly significant. Each column of Table 8 represents four Granger causality tests from other commodities towards the commodities in the first row for one direction. From Table 8, it is clear that the cattle and corn prices have little Granger prediction power for gold futures prices. However, there is significant indication that they can be driven by other commodities. On the other hand, copper and oil prices have Granger causality for gold futures prices. This shows that metal and oil are dominant forces in driving the overall commodity market, while agricultural commodities have little influence on other commodity prices. Therefore, we redo the gold cointegration test by removing cattle and corn prices since they present little prediction power for gold prices. Additionally, it might also be arguable that the nonstationarity of residuals stems from no time lags. As a result, we redo the two cointegration tests for different lags and check the stationarity of residuals. For the one-period time lag, both cattle and gold residuals still present nonstationarity. However, residuals for corn futures turn out to be stationary (see Table 9). On the other hand, when the regression is conducted for the two-period time lag, all three residuals become nonstationary (see Table 10). It is surprising that most residuals are still not stationary. It is subtle that even though they have Granger causality for each other, they are not integrated even for controlled time lags. It is believable that time lag may not be the crucial factor contributing to the nonstationarity of the residual. As a consequence, we propose that there should be a factor that interferes with the residuals and makes residuals nonstationary. We try to find the interference factor and filter it from the regression residuals such that the residuals can become stationary.where i = 1, 2, representing ca, cor, respectively.

$$\ln P_t^g = \alpha_g + \sum_{j=1, j\neq g, j\neq ca, j\neq cor}^{5} \beta_g \ln P_{t-1}^j + \varepsilon_{g,t}$$

This table reports the Dicky-Fuller test for residuals from the regression model with one lag of the commodity price, and the residuals for both gold and cattle price regression exhibit nonstationarity. However, the residual for corn price regression exhibits stationarity. Significance levels are denoted by *, **, ***, which correspond to the 10%, 5%, and 1% levels, respectively, and p-values are reported in brackets.

$$\ln P_t^i = \alpha_i + \sum_{j=1, j\neq i}^{5} \beta^j \ln P_{t-2}^j + \varepsilon_{i,t}$$

where i = 1, 2, representing ca, cor, respectively.

$$\ln P_t^g = \alpha_g + \sum_{j=1, j\neq g, j\neq ca, j\neq cor}^{5} \beta_g \ln P_{t-2}^j + \varepsilon_{g,t}$$

This table reports the Dicky-Fuller test for residuals from the regression model with two lags of the commodity price, and the residuals for all three commodities, cattle, corn and gold prices regression exhibit nonstationarity. Significance levels are denoted by *, **, ***, which

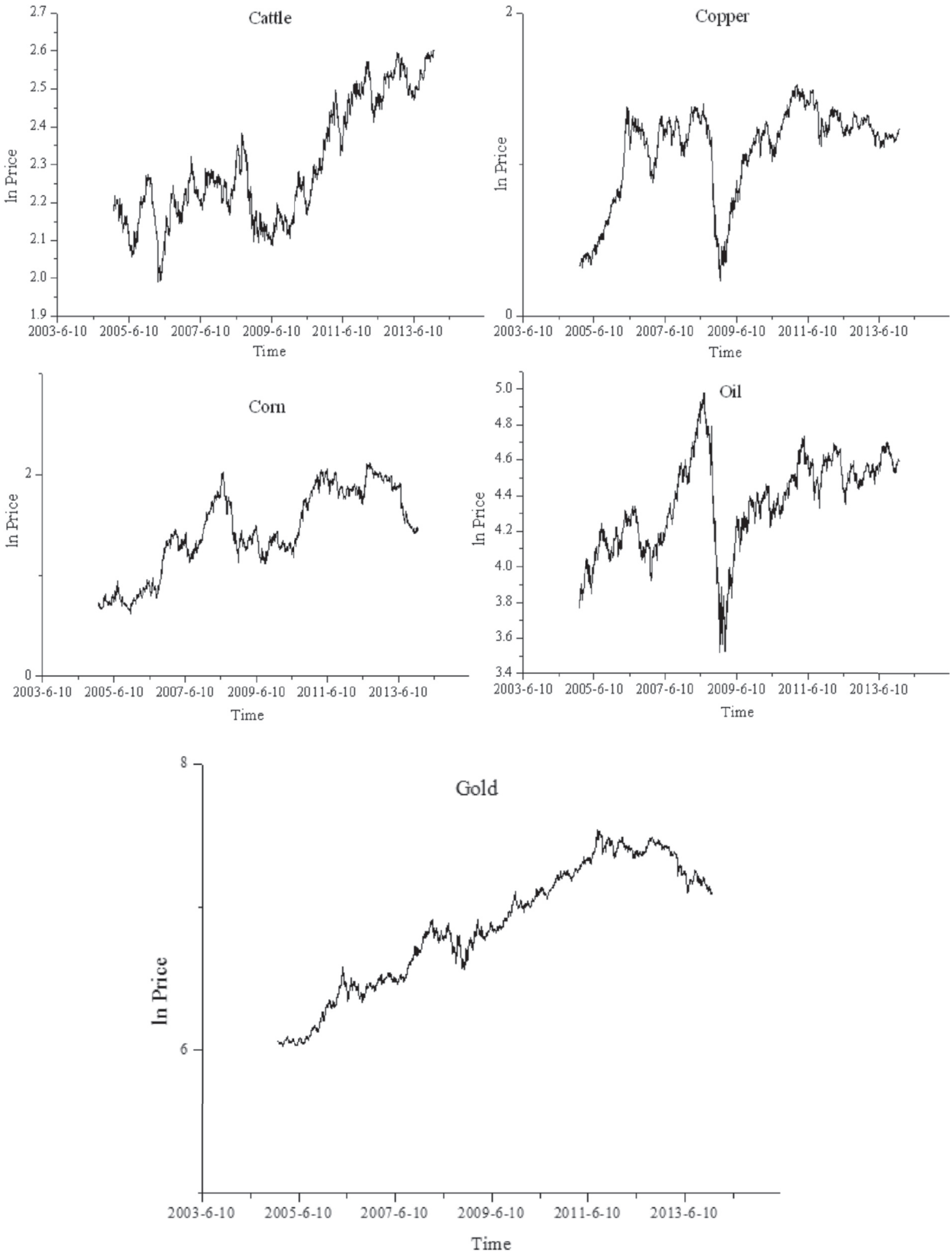

Fig. 1. Natural log of commodity daily prices.

correspond to the 10%, 5%, and 1% levels, respectively.

*It has been well documented that the liquidity level can contribute to commodity price comovements (*Zhang and Ding, 2018*;* Zhang et al., 2018*). Therefore, it is arguable that liquidity can be the* interference factor, where liquidity level can also *contribute to commodity price cointegration*. As a result, we control all the relevant liquidity variables from the regression equation to see whether the residuals can become stationary. Table 11 compares the two sets of results. The first set of results does not take liquidity into account, and some residuals are nonstationary. On the other hand, the second set of results takes liquidity as the control

Table 6
Regression results for commodity prices with other commodity markets. Cattle, corn and copper have both positive and negative impacts on other commodity prices, while gold and oil have positive impacts on most commodity prices. Significance levels are denoted by *, **, ***, which correspond to the 10%, 5%, and 1% levels, respectively, and standard errors are reported in brackets.

IV / DV	lnP^{ca}_t	lnP^{cor}_t	lnP^{cop}_t	lnP^{g}_t	lnP^{o}_t
lnP^{ca}_t	–	−0.084***	−0.17***	0.51***	0.0002
	–	(0.014)	(0.015)	(0.012)	(0.015)
lnP^{cor}_t	−0.13***	–	−0.025	0.72***	0.46***
	(0.023)	–	(0.019)	(0.013)	(0.016)
lnP^{cop}_t	−0.24***	−0.022	–	0.33***	0.66***
	(0.022)	(0.017)	–	(0.019)	(0.016)
lnP^{g}_t	0.73***	0.29***	0.70***	–	−0.33***
	(0.017)	(0.016)	(0.013)	–	(0.017)
lnP^{o}_t	0.0002	0.58***	0.46***	−0.32***	–
	(0.022)	(0.013)	(0.016)	(0.017)	–
$Adj\text{-}R^2$	0.51	0.66	0.77	0.81	0.67

Table 7
Dicky-Fuller test for residuals of the regression model.

	$\varepsilon_{ca,t}$	$\varepsilon_{cor,t}$	$\varepsilon_{cop,t}$	$\varepsilon_{g,t}$	$\varepsilon_{o,t}$
DF	−1.70	−2.68*	−3.74***	−2.45	−4.93***
	(0.42)	(0.08)	(0.00)	(0.12)	(0.00)

Table 8
Granger test results of the F-statistic of three commodities (with P-values in brackets); most commodities exhibit Granger causality for each other.

$$\ln P^i_t = \alpha_i + \sum_{j=1, j\neq i}^{5} \beta^j \ln P^j_{t-1} + \varepsilon_{i,t}$$

	lnP^{ca}_t	lnP^{cor}_t	lnP^{g}_t
lnP^{ca}_t	–	2.09*	1.02
	–	(0.08)	(0.39)
lnP^{cor}_t	9.15***	–	0.57
	(0.00)	–	(0.63)
lnP^{cop}_t	5.57***	2.03*	2.79*
	(0.01)	(0.09)	(0.06)
lnP^{g}_t	4.88**	2.38**	–
	(0.02)	(0.05)	–
lnP^{o}_t	11.46***	2.31*	4.31**
	(0.00)	(0.09)	(0.03)

Table 9
Dicky-Fuller test for residuals of the regression model.

	$\varepsilon_{ca,t}$	$\varepsilon_{cor,t}$	$\varepsilon_{g,t}$
DF	−2.02	−3.41***	−1.48
	(0.27)	(0.01)	(0.53)

Table 10
Dicky-Fuller test for residuals of the regression model.

	$\varepsilon_{ca,t}$	$\varepsilon_{cor,t}$	$\varepsilon_{g,t}$
DF	−1.22	−2.50	−0.74
	(0.67)	(0.11)	(0.83)

variable, and all residuals become stationary. We can maintain that most of the nonstationary part of the residuals stem from commodity futures liquidity. After we extract the liquidity information from the residuals, they become more stationary, and markets are cointegrated.

Table 11
Dicky-Fuller test for residuals of the regression model.

	$\varepsilon_{ca,t}$	$\varepsilon_{cor,t}$	$\varepsilon_{cop,t}$	$\varepsilon_{g,t}$	$\varepsilon_{o,t}$
DF without liquidity	−1.70	−2.68*	−3.74***	−2.45	−4.93***
	(0.42)	(0.08)	(0.00)	(0.12)	(0.00)
DF with liquidity	−3.41***	−3.51***	–	−2.76*	–
	(0.01)	(0.01)	–	(0.07)	–

$$\ln P^g_t = \alpha'_g + \sum_{j=1, j\neq g, j\neq ca, j\neq cor}^{5} \beta^j_g \ln P^j_t + \sum_{j=1, j\neq g, j\neq ca, j\neq cor}^{5} \gamma^j_g L^j_t + \varepsilon'_{g,t} \ln P^i_t$$
$$= \alpha'_i + \sum_{j=1, j\neq i}^{5} \beta^j_t \ln P^j_t + \sum_{j=1, j\neq i}^{5} \gamma^j_t L^j_t + \varepsilon'_{i,t} \tag{5}$$

where i = 1, 2, representing ca, cor, respectively.

This table reports the Dicky-Fuller test for residuals from the regression model. This regression model (equation (5)) has controlled the liquidity variables compared with the previous model (equation (4)). Then, we compared two test results from two models, where one model controls the liquidity and the other does not. From the p-values indicated in the brackets for the model that does not control liquidity, both the residuals of cattle and corn are not stationary, corn is weakly stationary and crude oil and copper futures are strongly stationary. On the other hand, the nonstationary residuals from the previous model all become stationary in the new model with liquidity, and all five residuals present stationarity. Significance levels are denoted by *, **, ***, which correspond to the 10%, 5%, and 1% levels, respectively.

The cointegration relation reveals the long-run equilibrium relation among different commodity prices. We have demonstrated that such long-run price equilibrium exists only for oil and copper futures. For the other three markets, gold, cattle and corn futures, liquidity interferes with the long-run cross-market commodity price equilibrium. After controlling the liquidity variables, gold, cattle and corn futures are cointegrated with other markets.

4. Aggregate liquidity effect on market price cointegration

In section 3, we have shown that individual price dynamics are driven by the coupling effect of cross-sectional price and liquidity evolution from other markets through cointegration analysis. In this section, we investigate whether the individual price evolution can also be explained by the aggregated market move. First, we test whether individual price is cointegrated with the aggregate market, namely, the CRB index. We conduct the following five regressions:

$$\ln P^i_t = \alpha_{1,i} + \alpha_{2,i} \ln P^{CRB}_t + \varepsilon^i_t \tag{6}$$

where $\mathrm{P}^{\mathrm{CRB}}_t$ is the CRB index and $\mathrm{P}^{\mathrm{i}}_t$ are individual prices, for i = 1, 2, 3, 4, 5 representing ca, cor, cop, g, o, respectively.

We test the stationarity of regression residuals for five commodities in equation (6). The results turn to be negative, where the residuals for most commodity futures prices are not stationary. Specifically, the Dicky-Fuller stationary test for residual ε^i_t is summarized in Table 12, from which we find that both copper and oil are cointegrated with the market CRB index, while the agriculture markets (cattle and corn) and gold market are not cointegrated with the overall CRB market index. The cointegration of copper and oil with the market index reveals that the metal and oil markets dominate the overall market CRB index. The noncointegration of cattle, corn and gold with the CRB index shows that agricultural and gold markets are relatively isolated from the global commodity index. The above interesting results may be explained by the different characteristics of commodity markets. Metal and oil markets are strongly influenced by the global economy; thus, they could be more likely predicted by global economic indicators, such as the CRB index. On the other hand, because agricultural prices are heavily regulated by local

Table 12
Dicky-Fuller test for five commodity futures with the CRB market index.

	$\varepsilon_{ca,t}$	$\varepsilon_{cor,t}$	$\varepsilon_{cop,t}$	$\varepsilon_{g,t}$	$\varepsilon_{o,t}$
DF	−1.58	−2.12	−2.26*	−0.75	−3.84***
	(0.49)	(0.13)	(0.08)	(0.83)	(0.00)

government due to their consumption purpose, their prices are less likely to be driven by the global CRB index level. Gold is a unique market that is mainly for investment and hedging purposes. Thus, gold is less relevant to the global commodity index.

This table reports the Dicky-Fuller test for residuals from the previous regression models (equation (6)). From the p-values indicated in the brackets, most of the residuals are nonstationary. Significance levels are denoted by *, **, ***, which correspond to the 10%, 5%, and 1% levels, respectively, and p-values are reported in brackets.

In the next step, we investigate whether the predictability of agricultural and gold markets can be improved by adding market liquidity information. Following Chordia et al. (2000)*, we use the equally weighted average liquidity measure as the indicator of the aggregate market liquidity measure to represent the liquidity commonality, which has been defined in* section 2*. We regress all five commodity future prices to the 21-day rolling averaged market liquidity measure (MLt-1) with control of the CRB market index.*

$$\ln P_t^{ca} = \alpha_1 + \beta_{11} ML_{t-1} + \beta_{12} \ln P_t^{CRB} + \varepsilon_t^{ca}$$
$$\ln P_t^{cop} = \alpha_2 + \beta_{21} ML_{t-1} + \beta_{22} \ln P_t^{CRB} + \varepsilon_t^{cop}$$
$$\ln P_t^{cor} = \alpha_3 + \beta_{31} ML_{t-1} + \beta_{32} \ln P_t^{CRB} + \varepsilon_t^{cor}$$
$$\ln P_t^{g} = \alpha_4 + \beta_{41} ML_{t-1} + \beta_{42} \ln P_t^{CRB} + \varepsilon_t^{g}$$
$$\ln P_t^{o} = \alpha_5 + \beta_{51} ML_{t-1} + \beta_{52} \ln P_t^{CRB} + \varepsilon_t^{o}$$

This table reports the Dicky-Fuller test for residuals from the previous regression model. From the p-values indicated in the brackets, most of the residuals are stationary except the residuals for cattle and gold futures. Significance levels are denoted by *, **, ***, which correspond to the 10%, 5%, and 1% levels, respectively, and standard errors are reported in brackets.

Table 13 shows that the residuals for cattle and gold regressions are still nonstationary, whereas the residual for the corn market becomes stationary after adding market liquidity. Cattle and gold prices tend to overreact to market information, so they may not be synchronous with the market index movement (Chen, 1998). Therefore, it might be difficult for them to maintain a long-run equilibrium relationship with market index movement, and thus, their noncointegration relation with market index may not result from the inference of market liquidity. Comparing Table 11 with Table 13, we found that all residuals are stationary in Table 11, whereas two of five residuals are nonstationary in Table 13. Consequently, as a good predictor, aggregate market liquidity might not be as effective as cross-sectional individual liquidity. The reason for the weakness of overall market liquidity in predicting individual prices is because aggregate market liquidity ignores the interaction effect and heterogeneous nature across individual liquidity dynamics.

From the above analysis, it may be arguable that the index could be dominated by oil because oil holds a heavy weight of the index (23% weight of the index). Thus, it would be interesting to see whether other commodity prices are cointegrated with oil and oil liquidity. Therefore, we construct equation (7) and conduct the regressions based on equation (7), and the regression results are reported in Table 14. Residuals from Table 14 have been tested in Dicky-Fuller with results reported in Table 15. Different markets exhibit heterogeneous behavior towards oil price and oil liquidity. It is observable that residuals for corn and copper prices are stationary, suggesting that corn and copper are firmly cointegrated with oil price and oil market liquidity. However, cattle price and gold price are not cointegrated with oil price and oil market liquidity. Since the cointegration method is often adopted in determining the long-run equilibrium relationship and then predicting future prices, the cattle

Table 13
Dicky-Fuller test for residuals from the regression model.

	$\varepsilon_{ca,t}$	$\varepsilon_{cor,t}$	$\varepsilon_{cop,t}$	$\varepsilon_{g,t}$	$\varepsilon_{o,t}$
DF	−1.65	−2.66*	−2.86**	−2.23	−3.74***
	(0.45)	(0.08)	(0.05)	(0.19)	(0.00)

Table 14
Regression results for commodity prices with oil price and oil liquidity. Significance levels are denoted by *, **, ***, which correspond to the 10%, 5%, and 1% levels, respectively, and standard errors are reported in brackets.

IV / DV	$lnP^{ca}{}_t$	$lnP^{cor}{}_t$	$lnP^{cop}{}_t$	$lnP^{g}{}_t$
$\boldsymbol{lnP^o{}_t}$	0.03*	0.72***	0.67***	0.39***
	(0.09)	(0.00)	(0.00)	(0.00)
$\boldsymbol{L^o{}_t}$	−0.09***	−0.09***	−0.06***	−0.19***
	(0.00)	(0.00)	(0.00)	(0.00)
$\boldsymbol{Adj\text{-}R^2}$	0.09	0.45	0.60	0.36

Table 15
Dicky-Fuller test for residuals of the regression model equation (7).

	$\varepsilon_{ca,t}$	$\varepsilon_{cor,t}$	$\varepsilon_{cop,t}$	$\varepsilon_{g,t}$
DF	−1.91	−2.77*	−3.35***	−2.07
	(0.32)	(0.06)	(0.01)	(2.05)

and gold prices cannot be predicted by using the oil price and oil market liquidity. More importantly, this result is almost identical to that in Table 13, suggesting that the cointegration relation with the commodity index is the same as that with oil and oil liquidity. This result might be attributed to the heavy weight oil possesses in the commodity index, and it can shed new insights for future oil price studies, such as oil price and other commodity price investigations (see Pal et al., 2019*).*

$$\ln P_t^{i} = \alpha_i + \gamma_1 \ln P_t^{o} + \gamma_2 \overline{L}_t^{o} + \varepsilon_{i,t} \quad (7)$$

where i = 1, 2, 3, 4, representing ca, cor, cop, g, respectively.

This table reports the Dicky-Fuller test for residuals from the regression model. From the p-values indicated in the brackets, all residuals are stationary at the 5% level. Significance levels are denoted by *, **, ***, which correspond to the 10%, 5%, and 1% levels, respectively, and p-values are reported in brackets.

5. Conclusions

In this paper, we have studied the liquidity effects on commodity price *cointegration with both cross-market individual prices and the global aggregate market index. Our main results can be summarized as follows:*

(1) *Metal and oil markets are cointegrated with both other commodity markets and the global commodity index. They are responsive to shocks from other commodity markets and can be predicted by cross-market commodity prices. Therefore, policy makers could use policies on other markets, such as the* agricultural market, to neutralize the external shock from *metal and oil markets.*
(2) Agricultural and gold markets are *cointegrated with other cross-market commodities only when other cross-market liquidities are controlled, so they can be predicted by using both commodity prices and liquidity levels in cross-markets.* Agricultural and gold markets are not *cointegrated with the global commodity index regardless of whether aggregated market liquidity is controlled. Since those markets are relatively segmented markets, they might not be quite responsive to shocks from other commodity markets. As a result, demand and supply*

for those markets may play vital roles, which policy makers could adopt for agricultural and gold price adjustment.

(3) Aggregated market liquidity is not as effective as cross-market individual liquidities in explaining individual commodity price dynamics.

(4) There are heterogeneous dynamic price adjustment processes among different commodity markets regarding oil prices and oil liquidity. Corn and copper prices are cointegrated with oil price and oil liquidity, whereas cattle and gold markets are not that sensitive to oil price and oil liquidity.

*Our research findings exhibit theoretical significance. Our results reveal the role of liquidity in capturing missing information that is not incorporated in price dynamics in an incomplete or inefficient market. According to Fama's Efficient Market Hypothesis (*Fama, 1965, 1970bib_Fama_1970*), price movement fully reflects all relevant information in the efficient market. As a result, metal and oil markets are near-efficient markets, and prices fully reflect information, which explains their cointegration with other cross-markets without incorporating liquidity information. Thus, oil and metal prices can be predicted using cross-market commodity prices. On the other hand,* agricultural and gold markets are segregated or inefficient markets. Therefore, their prices cannot fully capture all available information. Our significant theoretical discovery is that liquidity can account for missing information that is not embedded in price dynamics. Furthermore, liquidity plays more dominant roles than price in driving cross market commodity daily price movement in inefficient or incomplete markets, such as agricultural and gold markets, which tend to be illiquid and for which prices can be predicted only if we include both cross market prices and liquidity levels.

The practical implications of our research have several aspects. Since metal and oil markets are integrated with both cross-market prices and the global commodity index, traders can effectively hedge metal and oil price risk by taking appropriate positions in cross-market commodities or the global index. However, such cross-market or global market hedging would be difficult to apply to agricultural and gold markets due to their market incompleteness resulting from liquidity interference. To policy makers, to control inflation risk, both commodity prices and their liquidity need to be closely monitored. Cointegration analysis provides a reliable prediction framework of future commodity price movements by controlling liquidity information. Furthermore, we also demonstrate that *different commodity markets have different sensitivities to oil prices and oil liquidity, which sheds new insight for future research on the effects of oil price dynamics.*

Conflict of interest statement for the paper

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The work is sponsored by K.C.Wong Magna Fund in Ningbo University.

We are also very grateful for editor Prof. Sushanta Mallick and anonymous referees' help on this paper.

References

Alquist, R., Bhattarai, S., Coibion, O., 2019. Commodity-price comovement and global economic activity. J. Monetary Econ. forthcoming.
Berenguer-Rico, V., Gonzalo, J., 2014. Summability of stochastic processes—a generalization of integration for non-linear processes. J. Econom. 178, 331–341.
Chen, H., 1998. Price limits, overreactions, and price resolution in futures markets. J. Futures Mark. 18, 243–263.
Chinn, M.D., Coibion, O., 2014. The predictive content of commodity futures. J. Futures Mark. 34, 607–636.
Chordia, T., Roll, R., Subrahmanyam, A., 2000. Commonality in liquidity. J. Financ. Econ. 56, 3–28.
Chordia, T., Roll, R., Subrahmanyam, A., 2008. Liquidity and market efficiency. J. Financ. Econ. 87, 249–268.
Cody, B.J., L.O., Mills, 1991. The role of commodity prices in formulating monetary policy. Rev. Econ. Stat. 73, 358–365.
Coleman, M., 1990. Cointegration-based tests of daily foreign exchange market efficiency. Econ. Lett. 32, 53–59.
Corwin, S.A., Schultz, P., 2012. A simple way to estimate bid-ask spreads from daily high and low prices. J. Finance 67, 719–760.
Engle, R.F., C.W.J., Granger, 1987. Co-integration and error correction: representation, estimation and testing. Econometrica 35, 251–276.
Engsted, T., Gonzalo, J., Haldrup, N., 1997. Testing for multicointegration. Econ. Lett. 56, 259–266.
Fama, E.F., 1965. Random walks in stock market prices. Financ. Anal. J. 21, 55–59.
Fama, E.F., 1970. Efficient capital markets: a review of theory and empirical work. J. Finance 25, 383–417.
Frino, A., Mollica, V., Zhou, Z., 2014. Commonality in liquidity across international borders: evidence from futures markets. J. Futures Mark. 34, 807–818.
Garner, C.A., 1989. Commodity prices: policy target or information variable? Note, J. Money, Credit and Bank. 21, 508–514.
Goyenko, R., Holden, C., Trzcinka, C., 2009. Do liquidity measures measure liquidity? J. Financ. Econ. 92, 153–181.
Granger, C.W., 1986. Developments in the study of cointegrated economic variables. Oxf. Bull. Econ. Stat. 48, 213–228.
Hassan, A.M.H., 2003. Financial integration of stock markets in the Gulf: A multivariate cointegration analysis. Int. J. Bus. 8, 1–12.
Holtemöller, O., Mallick, S., 2016. Global food prices and monetary policy in an emerging market economy: the case of India. J. Asian Econ. 46, 56–70.
Iwatsubo, K., Watkins, C., Xu, T., 2018. Intraday seasonality in efficiency, liquidity, volatility and volume: platinum and gold futures in Tokyo and New York. J. Commod. Mark. 11, 59–71.
Korajczyk, R., Sadka, R., 2008. Pricing the commonality across alternative measures of liquidity. J. Financ. Econ. 87, 45–72.
Koski, J.L., Michaely, R., 2000. Prices, liquidity, and the information content of trades. Rev. Financ. Stud. 13, 659–696.
Kowalski, C., 2014. The most actively traded commodities [Online], Avaliable at. http://commodities.about.com/od/researchcommodities/a/most-liquid-commodity-markets.htm. (Accessed 12 July 2014).
Lardic, S., Mignon, V., 2008. Oil prices and economic activity: an asymmetric cointegration approach. Energy Econ. 30, 847–855.
Liu, W., 2006. A liquidity augmented capital asset pricing model. J. Financ. Econ. 82, 631–671.
Lof, M., Nyberg, H., 2017. Noncausality and the commodity currency hypothesis. Energy Econ. 65, 424–433.
Marshall, B., Nguyen, N., Visaltanachoti, N., 2013. Liquidity commonality in commodities. J. Bank. Finance 37, 11–20.
Mallick, S.K., Sousa, R.M., 2013. Commodity prices, inflationary pressures, and monetary policy: evidence from BRICS economies. Open Econ. Rev. 24 (4), 677–694.
N, ASDAQ, 2013. Commodity fact sheet [online], [Online], Avaliable at: http://www.nasdaqtrader.com/content/ProductsServices/dataproducts/RealTimeIndexes/NASDAQCommodityFactSheet.pdf. (Accessed 1 November 2016).
Pal, D., S.K., Mitra, 2019. Correlation dynamics of crude oil with agricultural commodities: a comparison between energy and food crops. Econ. Modell. 82, 453–466.
Roll, R., 1984. A simple implicit measure of the effective bid-ask spread in an efficient market. J. Finance 39, 1127–1139.
Serletis, A., 1994. A cointegration analysis of petroleum futures prices. Energy Econ. 16, 93–97.
Sockin, M., Xiong, W., 2015. Informational frictions and commodity markets. J. Finance 70, 2063–2098.
Westgaard, S.,M., Estenstad, M., Seim, Frydenberg, S., 2011. Co-integration of ICE Gas oil and Crude oil futures. Energy Econ. 33, 311–320.
Zhang, Y.J., Y.M., Wei, 2010. The crude oil market and the gold market: evidence for cointegration, causality and price discovery. Resour. Pol. 35, 168–177.
Zhang, Y., Ding, S., 2018. The return and volatility co-movement in commodity futures markets: the effects of liquidity risk. Quant. Finance 1471–1486.
Zhang, Y., Ding, S., Scheffel, E., 2018. Policy impact on volatility dynamics in commodity futures markets: evidence from China. J. Futures Mark. 38, 1227–1245.
Zhang, Y., Ding, S., Scheffel, E., 2019. A key determinant of commodity price Co-movement: the role of daily market liquidity. Econ. Modell. 81, 170–180.