

Contributions of Crude Oil Exchange Traded Funds in Price Discovery Process

American Business Review
Nov. 2020, Vol.23(2) 393 - 407
© The Authors 2020, CC BY-NC
ISSN: 2689-8810 (Online)
ISSN: 0743-2348 (Print)

Keshab Shrestha ^a, Sheena Philip ^a, and Yessy Peranginangin ^a

<https://doi.org/10.37625/abr.23.2.393-407>

ABSTRACT

This study empirically investigates the contributions of three crude oil-based exchange-traded funds (ETFs) in the price discovery process. Using daily data on the crude oil spot, near month crude oil futures, and three crude-oil-based ETFs, we analyze the price discovery contributions of the five-price series. We use two information share measures, namely the generalized information share (GIS) measure (Lien and Shrestha, 2014) and the permanent-temporary decomposition (PT/GG) measure (Gonzalo and Granger, 1995). We find that the futures market dominates the price discovery process. However, we also find that the crude-oil-based ETFs significantly contribute to the price discovery process. Thus, we find that additional ETFs play a significant role in price discovery. Therefore, they are not redundant in terms of their price discovery contributions.

KEYWORDS

Price Discovery, Cointegration, Information Share, Crude Oil

INTRODUCTION

Prices play an essential role in a free market economy. They help the so-called invisible hands (the term coined by the famous economist Adam Smith) achieve an optimal allocation of resources (Lien and Shrestha, 2009). However, to attain such allocation efficiency, the prices should reflect their fundamental values without any distortions. When the prices do not reflect their fundamental values, it will lead to market failure, and the free-market will not achieve the allocation efficiency. In such situations, regulatory interventions are usually required. Therefore, whether prices reflect the fundamental values is one of the critical questions in financial economics.

In analyzing whether prices reflect the fundamental values, it is essential to understand how the fundamental or true values get reflected in the prices. Price discovery is a process by which the new relevant information is incorporated into prices. As mentioned by O'Hara (O'Hara, 2003), in her presidential address to the American Finance Association, markets have two critical functions - liquidity and price discovery. Therefore, markets play an essential role in the price discovery process. In this study, we empirically analyze the contributions of three crude-oil-based exchange-traded funds (ETFs) to the price discovery process in the crude-oil-based markets.¹

^a Monash University, Malaysia

Corresponding Author:

Peranginangin (yessy.arnold@monash.edu)

We would like to thank Monash University Malaysia for the research support.

¹ In this study, we consider three oil-based ETFs together with crude oil spot and futures prices in the price discovery process. We neither analyze nor discuss the determinants of these prices. Our main focus is to empirically analyze the price discovery in the observed prices.

Oil-based ETFs are essentially managed funds that trade like shares (NYSE, 2017). These funds replicate the performance of crude oil through the acquisition of oil-related securities. The creation of oil-based ETFs lowers the barriers for financial commodity investors to participate in the crude oil market. Therefore, crude-oil-based ETFs could contribute to price discovery through several mechanisms. First, the ETFs have relatively lower holding and transaction costs that allow for smaller trades size and greater trade frequency (Murdock and Richie, 2008; Sukcharoen et al., 2015; Henderson et al., 2015; Ivanov, 2015). Moreover, Chu et al. (1999) suggest that the trading venue with lower transaction costs, lesser restrictions, higher liquidity, and greater access to leverage have the potential to lead the price discovery. These arguments are consistent with the empirical findings of Hasbrouck (2003), who finds that the price discovery takes place mostly in the electronically traded, small-denomination futures (E-mini) markets for the S&P500 and NASDAQ-100 stock indices. As for the S&P 500 index, where no E-mini contract existed over the sample period, the ETF market is found to provide substantial (approximately 50%) price discovery. Second, the shares of crude-oil-based ETFs are actively traded throughout the trading day that allows ETF shares to impound new information faster (Murdock and Richie, 2008; Henderson et al., 2015). Supporting this notion, Balchunas (2016) suggests that ETF prices provide real-time information that can be useful during normal and also when the markets are closed, which makes it a rival to the futures markets.

Prior studies have investigated the price discovery in various markets², including the crude-oil-based markets.³ The previous studies on the crude oil-based markets mainly examine the price discovery role of the futures and spot markets (Bopp and Sitzler, 1987; Bopp and Lady, 1991; Schwarz and Szakmary, 1994; Silvapulle and Moosa, 1999; Silverio and Szklo, 2012; Shrestha, 2014). The futures market is expected to dominate its spot counterpart in impounding new information due to its low transaction costs and greater flexibility for allowing short positions (Silverio and Szklo, 2012). On the other hand, Quan (1992) finds that the crude oil spot market dominates its futures counterparts in price discovery, but the two markets converge quite quickly. Bekiros and Diks (2008), Kaufmann and Ullman (2009), and Shrestha (2014) find that price discovery in crude oil takes place in both the spot and futures markets.

Ivanov (2013) investigates the price discovery role of one crude oil ETF, the United States Oil Fund (USO), using six months' worth of data in 2009, and finds that the USO contributes to the price discovery process. This study aims to investigate the price discovery contributions of two additional ETFs using a more extended period and more robust methodologies.⁴ More importantly, we would

² There exists an extensive number of studies on price discovery that investigate the share markets (e.g., Pascual et al., 2004; Kim et al., 2009; Riordan et al., 2013) and cross-listed shares (e.g., Hasbrouck, 1995; Harris et al., 1995; Lieberman et al., 1999; Eun and Sabherwal, 2003; Von Furstenberg and Tatora, 2004; Grammig et al., 2005; Pascual et al., 2006; Su and Chong, 2007; Chen et al., 2010; Frijns et al., 2010, 2015; Chen and Tourani-Rad, 2016). Some of the price discovery studies analyze price discovery in spot and derivative markets (e.g., Kawaller et al., 1987; Martikainen and Puttonen, 1994; Hasbrouck, 2003; Zhong et al., 2004; Lien and Shrestha, 2009; Rittler, 2012; Schultz and Swieringa, 2014; Shrestha, 2014; Kharbanda and Singh, 2017; Lin et al., 2018).

³ The studies on crude oil market have also looked into other aspects such as the relationship between crude oil and equity markets (e.g., Dagher and El Hariri, 2013; Cunado and Perez de Gracia, 2014; Ghosh and Kanjilal, 2016; Noor and Dutta, 2017; Dutta et al., 2017).

⁴ Ivanov (2013) uses Hasbrouck (1995) information share measure, which can only be applied to cases where the cointegrating relationships are one-to-one (Lien and Shrestha, 2009, 2014; Shrestha, 2014). Even though, there are theoretical reasons to justify a one-to-one cointegrating relationship between spot and futures prices, there is no such reason to expect that the cointegrating relationships between the spot and ETF would be one-to-one. Although, the cointegrating vector is not reported, looking at the log price difference reported in Ivanov (2013)'s Table 2, it is possible that the cointegrating vector is not one-to-one. This could be the reason that the monthly average IS measures for oil ETF vary from zero for August to maximum of around 91% for June (Ivanov (2013)'s Table 5). Similarly, the monthly average IS measures for oil futures contract vary from around 1% for June to around 98% for August.

like to determine if the additional ETFs play any price discovery role or if they are merely redundant in terms of price discovery.

This study uses two different price discovery measures: (i) the generalized information share (GIS) proposed by Lien and Shrestha (2014), and (ii) information share based on Gonzalo-Granger permanent-temporary (PT/GG) decomposition (Gonzalo and Granger, 1995). Both these measures do not require the cointegrating vectors to be one-to-one and, therefore, allowing more robust conclusions to be drawn with regards to the contributions of oil-based ETFs to the price discovery process.⁵

We use daily price data from December 2007 to December 2019 obtained from Datastream.⁶ The five price series used in this study include the crude oil spot (WTI spot Cushing), near month crude oil futures, and three ETFs USO (United States Oil Fund), OIL (iPath S&P GSCI Crude Oil) and USL (United States 12 Month Oil Fund).⁷ To estimate the price discovery or information share measures, we need to ensure that each of the five logarithms of price series considered in this study consists of a single unit-root. Furthermore, all five series should be cointegrated with four cointegrating vectors. To test for unit-root, we used the Phillips-Perron (PP) (Phillips and Perron, 1988) test, which incorporates unit-root as the null hypothesis. For the robustness, we also use the KPSS test (Kwiatkowski et al., 1992), where stationarity is the null hypothesis. Both tests indicate that each of the series has a single unit-root. As for the cointegration test, we use the Johansen (Johansen, 1991) cointegration test, which includes two tests: (i) λ max test and (ii) Trace test. Both tests indicate that there are four cointegrating vectors among the five-price series. However, some of the cointegrating relationships are far from being one-to-one. Therefore, the Hasbrouck information share (Hasbrouck, 1995) measure cannot be used to analyze the price discovery because this method requires the cointegrating vectors to be one-to-one. As mentioned above, in this study, we use GIS and PT/GG methods that do not require the cointegrating relationships to be one-to-one.

Based on GIS, we find that approximately 41 percent of the price discovery occurs in the futures market. We also find that approximately 25 percent of the price discovery takes place in the USO market. Finally, the USL and OIL markets contribute 11 and 6 percent, respectively. Overall, the three oil-based ETFs account for 42 percent of the price discovery. On the other hand, based on PT/GG information share measure, approximately 37 percent of price discovery takes place in the futures market. As to the ETFs, the USO and USL markets contribute about 37 and 14 percent, respectively. But, the third ETF (OIL) does not contribute to the price discovery. The total contribution of the three ETF markets is approximately equal to 50 percent. These results extend the findings of Ivanov (2013) through the inclusion of two additional oil-based ETFs that are found to contribute to the price discovery process. More importantly, we find that the additional ETFs play a significant price discovery

⁵ It is important to note that the econometric model used to measure the price discovery is based on sound financial economic theory and a well-accepted model. The basic idea is that efficient prices follow a class of processes called random walk or unit root process. If such unit root processes are cointegrated with the number of cointegrating vectors equal to the number of processes minus one, all these processes would be driven by a single unit root process known as the common stochastic trend (Stock and Watson, 1988). This common stochastic trend is considered to represent the fundamental value. The price discovery process measures the extent of the common stochastic trend reflected in various prices. The theories and models apply equally to security and commodity prices. This will be discussed in more detail in the methodology section.

⁶ We use daily instead of intraday price because the spot market does not seem to change too frequently during the day. The lack of movements in the intraday spot price would compromise the intraday price discovery analysis.

⁷ Our study only includes three oil-based ETFs that mimic the movements of WTI Cushing crude oil price by investing in the nearby month futures. We exclude oil ETFs that have share ownership in oil-related companies as well as ETFs that engage in leverage strategies. The exclusions allow us to capture a clean price discovery contribution of the oil ETFs that are not influenced by firm-specific or leverage risks.

role and are not redundant.⁸

We contribute to the literature in several ways. First, our findings provide comprehensive evidence on the price discovery role of three oil-based ETFs. Previous studies on price discovery in the crude oil market have not considered ETFs (Quan, 1992; Bekiros and Diks, 2008; Kaufmann and Ullman, 2009; Silverio and Szklo, 2012; Shrestha, 2014) with the exception of Ivanov (2013). Second, we extend Ivanov (2013) study on the price discovery role of ETFs by analyzing the role of the additional ETFs on price discovery. We also implement robust methodologies to estimate the information shares, including a more comprehensive sample of ETFs, and by including a longer sample period in the analysis. The additional ETFs are found to contribute to the price discovery process, and thus, they are not redundant in terms of price discovery contribution. Our empirical results also suggest that future studies on price discovery in the crude oil market would need to include ETFs into consideration. Third, our results allude to the discussion on the cost and benefit of the financialization of commodity markets. The significant price discovery role of oil-based ETFs may be considered as additional benefit of such financialization.

The remainder of this paper is as follows. The next section will discuss the methodology used in this paper, and Section 3 presents the empirical results. The study concludes in Section 4.

METHODOLOGY

In this section, we briefly discuss the methodologies used to analyze the price discovery contributions of the ETFs, where the price discovery is measured by information share (IS) measures. We use two methods to estimate the IS measures, namely (i) the generalized information share (GIS) proposed by Lien and Shrestha (2014), and (ii) the IS measure based on the Gonzalo-Granger permanent-temporary (PT/GG) decomposition (Gonzalo and Granger, 1995). First, let us look at the basic framework on which IS measures are based. Let Y_t be an $n \times 1$ vector of n unit-root or random walk series, where it is assumed that there are $(n-1)$ cointegrating vectors, that is, the system consists of a single common stochastic trend (Stock and Watson, 1988). The framework used in this study is based on well-established financial economic theories of efficient prices that apply to both security prices and commodity prices. For example, based on these theories developed by Samuelson (1965) and Fama (1965, 1970), efficient prices follow a random walk process. Several studies use the random walk as a model for efficient prices (e.g., Hasbrouck, 1991; Hasbrouck, 1993; Benink and Bossaerts, 2001; Boehmer et al., 2005; Boehmer and Kelley, 2009; Griffin et al., 2010; Campbell et al., 2011; Boehmer and Wu, 2013; Diebold and Strasser, 2013; Chaboud et al., 2014; Fotak et al., 2014; Boehmer et al., 2015; Conrad et al., 2015; Qin and Singal, 2015; Albuquerque et al. 2016; Lee et al., 2016; Brogaard et al., 2019; Hagstromer and Menkveld, 2019). Furthermore, we expect the prices considered in this study to be cointegrated because they are all related to the fundamental value of crude oil. The assumption of $(n-1)$ cointegrating vectors is based on our empirical finding.⁹

⁸ The price discovery contributions of OIL are approximately 6 and 0 percent based on GIS and PT/GG measures respectively. It is important to note that GIS is more reliable because the PT/GG method incorporates the way the prices react to new information where the nature of the information generation process is neglected, whereas the GIS method uses both the nature of information generation process as well as how the prices react to the new information (Lien and Shrestha, 2014). As a final note, we find the three out of five weights associated with the PT/GG methods have unexpected negative sign. This could be due to the fact that the underlying assumptions made by Gonzalo and Granger (1995) may not be valid in this case.

⁹ When we are considering n unit-root series, we cannot have n number of cointegrating vectors because this would imply the series to be stationary instead of unit-root series. If the number of cointegrating vectors is less than $(n-1)$, it would imply that there are more than one common stochastic trend (Stock and Watson, 1988). In such cases, the conventional IS measures cannot be applied. Since we found $(n-1)$ cointegrating vectors implying a single common stochastic trend which represents the fundamental value of the crude oil.

When the unit-root series are cointegrated, the vector autoregressive (VAR) model is misspecified. Therefore, the series have the following vector error-correction (VEC) representation (Engle and Granger, 1987), instead of VAR:

$$\Delta Y_t = \Pi Y_{t-1} + \sum_{i=1}^k A_i \Delta Y_{t-i} + \varepsilon_t, \quad \Pi = \alpha \beta^T \quad (1)$$

where β and α are $n \times (n - 1)$ matrices of rank $(n - 1)$. The columns of β consist of the $(n - 1)$ cointegrating vectors, and each column of α consists of the adjustment coefficients. The matrix Π is decomposed in such a way that $\beta^T Y_t$ represents the vector of $(n - 1)$ stationary series. Let Ω denote the $n \times n$ covariance matrix of the innovation vector, i.e., $E[\varepsilon_t \varepsilon_t^T] = \Omega$. Following Stock and Watson (1988), equation (1) can be transformed into the following two equivalent vector moving average (VMA) representations (Hasbrouck, 1995):

$$\Delta Y_t = \Psi(L) \varepsilon_t \quad (2)$$

$$Y_t = Y_0 + \Psi(1) \sum_{i=1}^t \varepsilon_i + \Psi^*(L) \varepsilon_t \quad (3)$$

Then, the Engle-Granger representation theorem (Engle and Granger, 1987) implies the following (De Jong, 2002 and Lehmann, 2002):

$$\beta^T \Psi(1) = 0 \quad \text{and} \quad \Psi(1) \alpha = 0 \quad (4)$$

Based on the above representations, $\Psi(1) \varepsilon_t$ represents the long-run impact of innovations on the unit-root series (Hasbrouck, 1995). Different information share measures considered by Hasbrouck (1995), Lien and Shrestha (2009), and Lien and Shrestha (2014) are based on this term.

GENERALIZED INFORMATION SHARE (GIS) MEASURES

Based on the above framework, we first discuss the GIS measure.¹⁰ Note that, in this study, we have five non-stationary series with four cointegrating vectors. Therefore, the cointegrating vector represented by matrix β can be expressed as follows:¹¹

$$\beta^T = \begin{bmatrix} 1 & -\gamma_1 & 0 & 0 & 0 \\ 1 & 0 & -\gamma_2 & 0 & 0 \\ 1 & 0 & 0 & -\gamma_3 & 0 \\ 1 & 0 & 0 & 0 & -\gamma_4 \end{bmatrix} \quad (5)$$

Let ψ_j^r be the j^{th} row of $\Psi(1)$. Then, $(n-1)$ cointegrating relations imply the following:

$$\psi_1^r = \gamma_{j-1} \psi_j^r, \quad j = 2, \dots, 5 \quad (6)$$

¹⁰ Please see Lien and Shrestha (2014) for detail.

¹¹ It is important to note that the Hasbrouck information share (Hasbrouck, 1995) requires all the cointegrating relationships to be one-to-one, i.e., $\gamma_i = 1.0$ for $i = 1, 2, \dots, 4$. The GIS method does not impose any such restrictions.

In other words, equation (4) implies that the first row of $\Psi(1)$ is equal to the γ_{j-1} times the j^{th} row of $\Psi(1)$. Therefore, the *long-run impact* of innovations on the i^{th} series is respectively given by

$$\psi_i^r \varepsilon_t = \psi_1^r \gamma_{i-1}^{-1} \varepsilon_t, \quad i = 1, \dots, 5 \quad (7)$$

where $\gamma_0 = 1$. When the innovations are *independent* (i.e., Ω is diagonal), the variances of *long-run impact* on the i^{th} series is given by:

$$\psi_i^r \Omega \psi_i^{rT} = \gamma_{i-1}^{-2} \sum_{j=1}^5 \psi_{1j}^2 \Omega_{jj} \quad (8)$$

where ψ_{ij} is the j^{th} element of ψ_i^r and Ω_{ij} is the $(i, j)^{th}$ element of Ω .

Let $S_{j,i}$ denote the contribution of the innovation of series j to the total variance of the *long-run impact* of innovation on series i . Then, we have the following

$$S_{j,i} = \frac{\psi_{1j}^2 \Omega_{jj}}{\psi_1^r \Omega \psi_1^{rT}} \quad (9)$$

Note that $S_{j,i}$ is independent of i . Therefore, the contribution of the innovation of series j to the total variance of the *long-run impact* of innovation on any series will be the same. It is important to note that $S_{j,i}$ given by equation (9) is valid only when the innovations are independent. However, in general, the innovations are not independent. When the innovations are not independent, we diagonalize the correlation matrix and end up with the GIS measure proposed by Lien and Shrestha (2014). Let Λ be a diagonal matrix containing the eigenvalues of the innovation correlation matrix on the diagonal, where the corresponding eigenvectors are given by the columns of matrix G . Then, we can calculate the information share of the j^{th} series as follows:

$$S_j^G = \frac{(\psi_j^G)^2}{\psi_1^r \Omega \psi_1^{rT}} \quad (10)$$

where $\psi^G = \psi_1^r F^M$, $F^M = [G\Lambda^{-1/2}G^T V^{-1}]^{-1}$ and ψ_j^G is the j^{th} element of ψ^G . The information share measure given by equation (10) is referred to as the generalized information share (GIS) measure. It can be shown that the GIS measure is independent of ordering. Therefore, the GIS method leads to a unique information share, unlike the upper and lower bound for Hasbrouck IS measure.¹²

GONZALO-GRANGER PERMANENT TEMPORARY DECOMPOSITION (PT/GG) INFORMATION SHARE

Here, we briefly describe the PT/GG method.¹³ Gonzalo and Granger (1995) propose a way of decomposing the vector of non-stationary series Y_t into permanent component f_t (non-stationary series) and transitory (stationary) component \hat{Y}_t , where the identification of these components is achieved by assuming that (i) the permanent component is a linear function of the original series and that (ii) the transitory component does not Granger cause the permanent component in the long-run.

¹² See Lien and Shrestha (2009) and Lien and Shrestha (2014) for detail on this issue.

¹³ See Booth et al. (1999), Baillie et al. (2002), Booth et al. (2002), Harris et al. (2002), Lien and Shrestha (2009) and Figuerola-Ferretti and Gonzalo (2010) for more information on this method. However, the PT/GG method is usually discussed in situations with only two unit-root series. In this study, we deal with 5 unit-root series. Therefore, some necessary modifications are needed.

The permanent component f_t (under the linearity condition) can be written as

$$f_t = \mu^T Y_t \quad (11)$$

where μ is an $n \times 1$ (or, 5×1 in this study) permanent component coefficient vector which can be shown to be orthogonal to the adjustment coefficient matrix α , i.e., $\mu = \alpha_{\perp}$.

In our study, since we have five unit-root series, the permanent component, f_t can be represented by

$$f_t = \mu_1 Y_{1t} + \mu_2 Y_{2t} + \dots + \mu_5 Y_{5t} \quad (12)$$

If all the elements of μ are non-negative, then the PT/GG based information share of the i^{th} series is given by

$$PTGG_i = \frac{\mu_i}{\sum_{j=1}^5 \mu_j} \quad (13)$$

In the case of non-negative μ , its elements correspond to the contribution of individual series to the permanent component. Therefore, the above definition of price discovery makes sense. However, there is no guarantee that all the elements of μ will be non-negative. It is important to note that the negative elements of μ may not necessarily lead to problems. For example, if some of the unit-root processes are negatively related to the common-stochastic trend, we expect the corresponding elements of μ to be negative. In such cases, where the negative elements of μ are valid, we can use the following definition of information share:¹⁴

$$PTGG_i^+ = \frac{|\mu_i|}{\sum_{j=1}^5 |\mu_j|} \quad (14)$$

However, if some of the elements of μ are unexpectedly negative, one way to compute the PT/GG based information share is to replace the negative elements with zero and use equation (13). However, this will lead to all negative elements having zero information share regardless of their absolute value. Alternatively, we can add a constant number to each element of μ to obtain μ^* so that the series with the most negative element in μ will have zero information share.¹⁵ This method will result in the following information share measure:

$$PTGG_i^* = \frac{\mu_i^*}{\sum_{j=1}^5 \mu_j^*}, \quad \mu_i^* = \mu_i + (-\min\{\mu_i\}) \quad (15)$$

EMPIRICAL RESULTS

The data starts from 6 December 2007 to 31 December 2019, with a total number of 3,012 observations. We chose the starting date to be 6 December 2007 because this the earliest date for which the data for all five prices are available. The data set includes the daily WTI crude oil spot and near month crude oil futures prices. The data set also includes prices of three crude oil future-based ETFs. We focus on oil ETFs that replicate the WTI spot Cushing through the nearby month futures contracts. The three oil ETFs that replicate the WTI spot Cushing through the nearby month futures contracts. The three ETFs are

¹⁴ For example, suppose the common-stochastic trend is the *fundamental* value of gold. In this case, the unit-root processes could be the stock prices of gold mining companies. One of the unit-root series could be the price of a put option on gold. In this case, the element of μ corresponding to the put option is expected to be negative, because is it negatively related to the fundamental value of gold.

¹⁵ This method will also preserve the orthogonality of μ^* to the adjustment coefficient matrix α .

the United States Oil Fund (USO), iPath S&P GSCI Crude Oil Total Return (OIL), and United States 12 Month Oil Fund (USL). Throughout the empirical analyses, we use the logarithm of the prices. Table 1 presents detailed information on the three oil-based ETFs, including their ticker codes, inception dates, asset under management, and trading volumes. The USO accounts for around 77% of the assets under management, while OIL and USL account for 19% and 4%, respectively.

Table 1. Information of ETFs

This table presents the ticker codes, inception dates, assets under management and trading volumes for three ETFs used in the study.

Ticker Code	Oil ETFs	Inception Date	Asset Under Management (in USD million)	Trading Volume
U:USO	United States Oil Fund	10-04-2006	2,510	16,012,332
U:OIL	iPath S&P GSCI Crude Oil Total Return ETN	16-08-2006	628.87	4,107,677
U:USL	United States 12 Month Oil Fund Limited Partnership	06-12-2007	104.84	26,720

The price discovery is measured using GIS and PT/GG information share measures. To compute these two IS measures, we need to validate that each of the log-price series used in this study is non-stationary with a single unit-root. We use the Phillips-Perron (PP) (Phillips and Perron, 1988) unit-root test, where the unit-root is the null hypothesis. Table 2 presents the unit-root test results for the crude oil spot, futures, and the three ETFs.

Table 2. Unit Root Test for Crude Oil Spot, Futures, and ETFs

This table summarizes the results of the Philips-Perron (PP) and Kwiatkowski-Philips-Schmidt-Shin (KPSS) unit root test statistics for the natural logarithm of the crude oil spot, futures, and ETFs prices. The critical values for PP test are -2.568, -2.863, and -3.436 at the 10%, 5%, and 1% significance levels, respectively. The critical values for the KPSS test are 0.347, 0.463, and 0.739 at the 10%, 5%, and 1% significance levels, respectively.

Series	Philip-Perron Test		KPSS Test	
	Log of Price	First Difference	Log of Price	First Difference
Spot	-2.1510	-58.9358***	3.5972***	0.0481
Futures	-2.1295	-58.052***	3.6311***	0.0487
USO	-1.3661	-56.7887***	8.5566***	0.0803
OIL	-1.3459	-56.7939***	8.6254***	0.0843
USL	-1.5251	-56.8049***	7.1965***	0.0655

***, **, and * indicate the test statistic to be significant at the 1%, 5%, and 10% significance level, respectively.

All the PP test statistics for log-price series are insignificant, even at the 10% level. Also, all the PP test statistics for the first differenced log-price series are significant even at the 1% level. These results indicate that all five series are integrated at order one, i.e., each of the series consists of a single unit-root. To check the robustness of our test results, we also apply the KPSS test (Kwiatkowski et al., 1992), where stationarity is the null hypothesis. As reported in Table 2, the KPSS test statistics for log-prices are highly significant even at the 1% level, which rejects the null hypothesis of stationarity.

However, none of the KPSS statistics is significant for the first differenced series even at the 10% level. Therefore, both PP and KPSS tests reveal a consistent result where each of the series consists of a single unit-root.

We then proceed to examine the number of cointegrating vectors. As discussed earlier, we are dealing with five series. Therefore, to compute GIS and PT/GG based IS measures, we need to establish that there are four cointegrating vectors. We apply the Johansen (1991) cointegration test to find the number of cointegrating vectors. Table 3 reports the λ_{max} and Trace statistics. The λ_{max} statistics for the number of cointegrating vectors (r) less than or equal to 3 is significant even at the 1% level.¹⁶ However, the λ_{max} statistics for the number of cointegrating vectors (r) less than or equal to 4 is insignificant even at the 10% level. Similarly, the Trace statistics for the number of cointegrating vectors (r) less than or equal to 3 is significant at the 1% level. Therefore, we conclude that the number of cointegrating vectors is equal to 4.

Table 3. Cointegration Test for Crude Oil Spot, Futures, and ETFs

This table summarizes the results of the Johansen tests used to determine the number of cointegrating vectors. The AIC criterion is used to determine the lag length.

	λ_{max}	Critical Values			Trace	Critical Values		
		10%	5%	1%		10%	5%	1%
$r \leq 4$	3.54	7.52	9.24	12.97	3.54	7.52	9.24	12.97
$r \leq 3$	25.46***	13.75	15.67	20.2	29***	17.85	19.96	24.60
$r \leq 2$	45.97***	19.77	22.00	26.81	74.97***	32.00	34.91	41.07
$r \leq 1$	176.39***	25.56	28.14	33.24	251.36***	49.65	53.12	60.16
$r = 0$	551.07***	31.66	34.40	39.79	802.43***	71.86	76.07	84.45

***, **, and * indicate the test statistic to be significant at the 1%, 5%, and 10% significance levels, respectively.

The estimated cointegrating matrix β^T is given by

$$\beta^T = \begin{bmatrix} 1 & -0.9999 & 0 & 0 & 0 \\ 1 & 0 & -0.4034 & 0 & 0 \\ 1 & 0 & 0 & -0.3467 & 0 \\ 1 & 0 & 0 & 0 & -0.6690 \end{bmatrix}$$

with $\gamma_1 = 0.9999$, $\gamma_2 = 0.4034$, $\gamma_3 = 0.3467$, $\gamma_4 = 0.6690$. Even though, γ_1 is approximately equal to 1, other γ 's are significantly different from 1.0. Therefore, Hasbrouck information shares cannot be computed in this case. The GIS and PT/GG methods are the right methods to use in such a case. The GIS measures for the five series are presented in Table 4. Only approximately 41 percent of the price discovery takes place in the futures market. The remaining 59 percent of the price discovery takes place in the other four markets that include the spot and the three ETF markets. Approximately 25 percent of the price discovery occurs in the USO market. As to the remaining two ETFs, around 11 and 6 percent of the price discovery take place in the USL and OIL markets, respectively. In total, approximately 42 percent of the price discovery takes place in the three ETFs markets. These results

¹⁶ Strictly speaking, the λ_{max} test statistic is a likelihood ratio test statistic with the null hypothesis of number of cointegrating vector equal to r against the number of cointegrating vector equal to $r + 1$.

indicate that ETFs have significant price discovery contribution. Furthermore, each ETF plays a significant role in price discovery even though they differ in the extent of their contributions. As to the PT/GG information share measure, the elements of μ are presented in Table 4, where three of the five elements of μ are negative. Therefore, PT/GG information shares are computed based on equation (15) and are reported in Table 4. Based on these measures, we can conclude that the futures and USO markets contribute equally, about 37 percent, to the price discovery process. On the other hand, OIL has no contribution. The total contribution of the three ETF markets is approximately equal to 50 percent.¹⁷ In sum, we find that ETFs play a significant price discovery role, and additional ETFs play an additional price discovery role. Therefore, additional ETFs are not redundant.

Table 4. Information Share

This table summarizes the Generalized Information Share (GIS) for the spot, futures, and three ETFs. It also presents the PT/GG based information share measure. Since three of the elements of μ are negative, PTGG* is computed based on equation (15)

	Generalized Information Share		
	GIS	$\mu = \alpha_{\perp}$	PTGG*
Spot	0.1732	-0.0439	0.1276
Futures	0.4045	0.6450	0.3686
USO	0.2453	0.6438	0.3682
OIL	0.0634	-0.4087	0.0000
USL	0.1136	-0.0211	0.1356

CONCLUSION

In this study, we empirically investigate the contributions of the crude oil futures market and three crude-oil-based ETFs in the price discovery process. We use daily data from December 2007 to December 2019. We find that the crude oil spot, futures, and three ETF prices have a single unit-root. We also find that there are four cointegrating relationships among the five-price series. However, some of the cointegrating relationships are significantly different from being one-to-one. To analyze the price discovery contribution of the ETF markets, we apply two recently developed information share measures, which include the generalized information share (GIS) and Gonzalo and Granger permanent-temporary decomposition (PT-GG) based information share measures. Based on the GIS, we find that around 41 percent of the price discovery occurs in the futures market, and about 42 percent of the price discovery taking place in the three ETF markets. Based on the PT/GG measures, we find that about 37 percent of the price discovery takes place in the futures market, and about 50 percent of the price discovery taking place in the three ETF markets. Therefore, we conclude that ETF markets play a significant role in the price discovery process. More importantly, the additional ETF markets significantly contribute to the price discovery process implying that these markets are not redundant in terms of their price discovery contributions. Therefore, the ETF markets help the free market economy move towards the optimal allocation of resources by contributing to the price discovery process. Given the information on the additional price discovery role played by the oil-based

¹⁷ It is interesting to note that, according to Shrestha's (2014) findings, 45% and 40% of price discovery, based respectively on the GIS and PT/GG measures, are found to take place in the spot market for crude oil where no ETF are included in the analysis. Our analysis shows that the inclusion of three ETFs takes the price discovery away mostly from the spot market. For example, the contribution from the spot market is only 17 percent based on the GIS measure. Finally, as mentioned in the Introduction, GIS method is more reliable compared to PT/GG method. The negative values for the three elements of μ cannot be explained.

ETFs, policymakers, and regulators may need to encourage the development of such markets based on oil as well as ETFs based on other commodities.

REFERENCES

- Albuquerque, R., Eichenbaum, M., Luo, V.X., & Rebelo, S. (2016). Valuation risk and asset pricing. *The Journal of Finance*, LXXI (6), 2861-2904.
- Baillie, R., Booth, G., Tse, Y., & Zobotina, T. (2002). Price discovery and common factor models. *Journal of Financial Markets*, 5, 309-321.
- Balchunas, E. (2016). *The institutional ETF toolbox: How institutions can understand and utilize the fast-growing world of ETFs*. John Wiley & Sons, Inc., Hoboken, New Jersey.
- Bekiros, S.D., & Diks, C.G.H. (2008). The relationship between crude oil spot and futures prices: Cointegration, linear and nonlinear causality. *Energy Economics*, 2673-2685.
- Benink, H., & Bossaerts, P. (2001). An exploration of Neo-Austrian Theory applied to financial markets. *The Journal of Finance*, 56(3), 1011-1027.
- Boehmer, E., Saar, G., & Yu, L. (2005). Lifting the veil: An analysis of pre-trade transparency at the NYSE. *The Journal of Finance*, 60(2), 783-815.
- Boehmer, E., & Kelley, E.K. (2009). Institutional investors and the informational efficiency of prices. *The Review of Financial Studies*, 22(9), 3563-3594.
- Boehmer, E., & Wu, J. (2013). Short selling and the price discovery process. *The Review of Financial Studies*, 26(2), 287-322.
- Boehmer, E., Chava, S., & Tookes, H.E. (2015). Related securities and equity market quality: The case of CDS. *The Journal of Financial and Quantitative Analysis*, 50(3), 509-541.
- Booth, G., Lin, J., Martikainen, T., & Tse, Y. (2002). Trading and pricing in upstairs and downstairs stock markets. *The Review of Financial Studies*, 15(4), 1111-1135.
- Booth, G., So, R., & Tse, Y. (1999). Price discovery in German equity index derivatives markets. *The Journal of Futures Markets*, 19(6), 619-643.
- Bopp, A. E., & Lady, G. M. (1991). A comparison of petroleum futures versus spot prices as predictors of prices in future. *Energy Economics*, 13(4), 274-282.
- Bopp, A. E., & Sitzler, S. (1987). Are petroleum futures prices good predictors of cash value? *The Journal of Futures Markets*, 7(6), 705-719.
- Brogaard, J., Hendershott, T., & Riordan, R. (2019). Price discovery without trading: Evidence from limit orders. *The Journal of Finance*, LXXIV (4), 1621-1658.
- Campbell, J.Y., Giglio, S., & Pathak, P. (2011). Forced sales and house prices. *The American Economic Review*, 101(5), 2108-2131.
- Chaboud, A.P., Chiquoine, B., Hjalmarsson, E., & Vega, C. (2014). Rise of the machines: Algorithmic trading in the foreign exchange market. *The Journal of Finance*, 69(5), 2045-2084.
- Chen, J., & Tourani-Rad, A. (2016). Short sales and price discovery of Chinese cross-listed firms. *International Journal of Managerial Finance*, 12(4), 408-421.
- Chen, K., Guangzhong, L., & Wu, L. (2010). Price discovery for segmented us-listed Chinese stocks: location or market quality? *Journal of Business Finance and Accounting*, 38, 242-269.
- Chu, Q.C., Hsieh, W.I.G., & Tse, Y. (1999). Price discovery on the S&P500 index markets: An analysis of spot index, index futures, and SPDRs. *International Review of Financial Analysis*, 8(1), 21-34.
- Conrad, J., Wahal, S., & Xiang, J. (2015). High-frequency quoting, trading, and the efficiency of prices. *Journal of Financial Economics*, 116, 271-291.
- Cunado, J., & Perez de Gracia, F. (2014). Oil price shocks and stock market returns: Evidence for some European countries. *Energy Economics*, 42, 365-377.
- Dagher, L., & El Hariri, S. (2013). The impact of global price shocks on the Lebanese stock market. *Energy*, 63, 366-374.

- De Jong, F. (2002). Measures of contributions to price discovery: A comparison. *Journal of Financial Markets*, 5, 323-327.
- Diebold, F.X., & Strasser, G. (2013). On the correlation structure of microstructure noise: A financial economic approach. *The Review of Financial Studies*, 80(4), 1304-1337.
- Dutta, P., Noor, M.H., & Dutta, A. (2017). Impact of oil volatility stocks on global emerging market stock returns. *International Journal of Managerial Finance*, 13(5), 578-591.
- Engle, R.F., & Granger, C.W.J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2), 251-276.
- Eun, S., & Sabherwal, S. (2003). Cross-border listing and price discovery: evidence from US-listed Canadian stocks. *Journal of Finance*, 58, 549-575.
- Fama, E.F. (1965). The behavior of stock-market prices. *The Journal of Business*, 38(1), 34-105.
- Fama, E.F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2), 383-417.
- Figuerola-Ferretti, I., & Gonzalo, J. (2010). Modelling and measuring price discovery in commodity markets. *Journal of Econometrics*, 158, 95-107.
- Fotak, V., Raman, V., & Yadav, P.K. (2014). Fails-to-deliver, short selling, and market quality. *Journal of Financial Economics*, 114, 493-516.
- Frijns, B., Gilbert, A., & Tourani-Rad, A. (2010). The dynamics of price discovery for cross listed shares: Evidence from Australia and New Zealand. *Journal of Banking & Finance*, 34, 498-508.
- Frijns, B., Gilbert, A., & Tourani-Rad, A. (2015). The determinants of price discovery: Evidence from US-Canadian cross-listed shares. *Journal of Banking & Finance*, 59, 457-468.
- Ghosh, S., & Kanjilal, K. (2016). Co-movement of international crude oil price and Indian stock market: Evidences from nonlinear cointegration tests. *Energy Economics*, 53, 111-117.
- Griffin, J.M., Kelly, P.J., & Nardari, F. (2010). Do market efficiency measures yield correct inferences? A comparison of developed and emerging markets. *The Review of Financial Studies*, 23(8), 3225-3277.
- Gonzalo, J., & Granger, C. (1995). Estimation of common long-memory components in cointegrated systems. *Journal of Business and Economic Statistics*, 13(1), 27-35.
- Grammig, J., Melvin, M., & Schlag, C. (2005). Internationally cross-listed stock prices during overlapping trading hours: price discovery and exchange rate effects. *Journal of Empirical Finance*, 12, 139-164.
- Hagstromer, B., & Menkveld, A.J. (2019). Information revelation in decentralized markets. *The Journal of Finance*, LXXIV (6), 2751-2787.
- Harris, F., McInish, T., Shoesmith, G., & Wood, R. (1995). Cointegration, error correction, and price discovery on three informationally-linked security markets. *Journal of Financial and Quantitative Analysis*, 30, 563-579.
- Harris, F., McInish, T., & Wood, R. (2002). Security price adjustment across exchanges: in investigation of common factor components for Dow stocks. *Journal of Financial Markets*, 5, 277-308.
- Hasbrouck, J. (1991). The summary informativeness of stock trades: An econometric analysis. *The Review of Financial Studies*, 4(3), 571-595.
- Hasbrouck, J. (1993). Assessing the quality of a security market: A new approach to transaction-cost measurement. *The Review of Financial Studies*, 6(1), 191-212.
- Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *The Journal of Finance*, 50, 1175-1199.
- Hasbrouck, J. (2003). Intraday price formation in U.S. equity index markets. *The Journal of Finance*, 58(6), 2375-2399.

- Henderson, B.J., Pearson, N.D., & Wang, L. (2015). New evidence on the financialization of commodity markets. *The Review of Financial Studies*, 28, 1285-1311.
- Ivanov, S.I. (2013). The influence of ETFs on the price discovery of gold, silver and oil. *Journal of Economics and Finance*, 37, 453-462.
- Ivanov, S.I. (2015). Intraday analysis of currency ETFs. *International Journal of Managerial Finance*, 11(4), 438-450.
- Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, 59(6), 1551-1580.
- Kaufmann, R.K., & Ullman, B. (2009). Oil prices, speculation, and fundamentals: Interpreting causal relations among spot and futures prices. *Energy Economics*, 31, 550-558.
- Kawaller, I., Koch, P., & Koch, T. (1987). The temporal price relationship between S&P 500 futures and the S&P 500 index. *Journal of Finance*, 42, 1309-1329.
- Kharbanda, V., & Singh, A. (2017). Lead-lag relationship between futures and spot FX market in India. *International Journal of Managerial Finance*, 5, 560-577.
- Kim, S., Kim, I.J., & Nam, S.O. (2009). The lead-lag relationship between stock index options and stock index market: Model, moneyness, and news. *International Journal of Managerial Finance*, 5(3), 311-332.
- Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54, 159-178.
- Lee, C., Chung, K.H., & Yang, S. (2016). Corporate governance and the informational efficiency of prices. *Financial Management*, 45(1), 239-260.
- Lehmann, B. (2002). Some desiderata for the measurement of price discovery across markets. *Journal of Financial Markets*, 5, 259-276.
- Lieberman, O., Ben-Zion, U., & Hauser, S. (1999). A characterization of the price behavior of international dual stocks: an error correction approach. *Journal of International Money and Finance*, 18, 289-304.
- Lien, D., & Shrestha, K. (2009). New information share measure. *The Journal of Futures Markets*, 29(4), 377-395.
- Lien, D., & Shrestha, K. (2014). Price discovery in interrelated markets. *The Journal of Futures Markets*, 34(3), 203-219.
- Lin, C.B., Chou, R.K., & Wang, G.H. (2018). Investor sentiment and price discovery: Evidence from the pricing dynamics between the futures and spot markets. *Journal of Banking & Finance*, 90, 17-31.
- Martikainen, T., & Puttonen, V. (1994). International price discovery in Finnish stock index futures and cash markets. *Journal of Banking & Finance*, 18, 809-822. Special Issue on European Derivative Markets.
- Murdock, M., & Richie, N. (2008). The united states oil fund as a hedging instrument. *Journal of Asset Management*, 9, 333-346.
- Noor, M.H., & Dutta, A. (2017). On the relationship between oil and equity markets: Evidence from South Asia. *International Journal of Managerial Finance*, 13(3), 287-303.
- NYSE (2017). Exchange traded funds. URL: <https://www.nyse.com/products/etp-funds-etf>.
- O'Hara, M. (2003). Presidential address: Liquidity and price discovery. *The Journal of Finance*, 58, 1335-1354.
- Pascual, R., Escribano, A., & Tapia, M. (2004). Adverse selection costs, trading activity and price discovery in the NYSE: An empirical analysis. *Journal of Banking & Finance*, 28, 107-128.

- Pascual, R., Pascual-Fuster, B., & Climent, F. (2006). Cross-listing, price discovery and the informativeness of the trading process. *Journal of Financial Markets*, 9, 144-161.
- Phillips, P.C.B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2), 335-346.
- Qin, N., & Singal, V. (2015). Indexing and stock price efficiency. *Financial Management*, 44(4), 875-904.
- Quan, J. (1992). Two-step testing procedure for price discovery role of futures prices. *The Journal of Futures Markets*, 12(2), 139-149.
- Riordan, R., Storkenmaier, A., Wagener, M., & Zhang, S.S. (2013). Public information arrival: Price discovery and liquidity in electronic limit order markets. *Journal of Banking & Finance*, 37, 1148-1159.
- Rittler, D. (2012). Price discovery and volatility spillovers in the European union emissions trading scheme: A high-frequency analysis. *Journal of Banking and Finance*, 36, 774-785.
- Samuelson, P.A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6, 41-49.
- Schultz, E., & Swieringa, J. (2014). Catalysts for price discovery in the European union emissions trading system. *Journal of Banking & Finance*, 42, 112-122.
- Schwarz, T.V., & Szakmary, A.C. (1994). Price discovery in petroleum market: Arbitrage, cointegration, and the time interval of analysis. *The Journal of Futures Markets*, 14(2), 147-167.
- Shrestha, K. (2014). Price discovery in energy markets. *Energy Economics*, 45, 229-233.
- Silvapulle, P., & Moosa, I.A. (1999). The relationship between spot and futures prices: Evidence from the crude oil market. *The Journal of Futures Markets*, 19(2), 175-193.
- Silverio, R., & Szklo, A. (2012). The effect of the financial sector on the evolution of oil prices: Analysis of the contribution of the futures market to the price discovery process in the WTI spot market. *Energy Economics*, 2012, 1799-1808.
- Stock, J.H., & Watson, M.W. (1988). Testing for common trends. *Journal of the American Statistical Association*, 83, 1097-1107.
- Su, Q., & Chong, T. (2007). Determining the contributions to price discovery for Chinese cross-listed stocks. *Pacific-Basin Finance Journal*, 15, 140-153.
- Sukcharoen, K., Choi, H., & Leatham, D.J. (2015). Optimal gasoline hedging strategies using futures contracts and exchange-traded funds. *Applied Economics*, 47(32), 3482-3498.
- Von Furstenberg, G., & Tatora, C. (2004). Bolsa or NYSE: price discovery for Mexican shares. *Journal of International Financial Markets Institutions and Money*, 14, 295-311.
- Zhong, M., Darrat, A.F., & Otero, R. (2004). Price discovery and volatility spillovers in index futures markets: Some evidence from Mexico. *Journal of Banking & Finance*, 28, 3037-3054.