Article

Price Discovery in Agricultural Markets

American Business Review
May 2020, Vol.23(1) 53 - 69

ISSN: 2689-8810 (Online)
ISSN: 0743-2348 (Print)

Keshab Shrestha[a*], Ravichandran Subramaniam[a], and Thangarajah Thiyagarajan[a]

https://doi.org/10.37625/abr.23.1.53-69

ABSTRACT

In this study, we empirically analyze the contribution of futures markets to the price discovery process for seven agricultural commodities using the generalized information share proposed by Lien and Shrestha (2014) and component share based on the permanent-temporary decomposition proposed by Gonzalo and Granger (1995). We find that most of the price discovery takes place in the futures markets with the exception of cocoa. Our results show that futures markets play an important role in price discovery process. These results are important to academicians, practitioners, policymakers as well as business leaders.

KEYWORDS

Cointegration, Information Share, Price Discovery

JEL Classifications: C3, G1, Q4

INTRODUCTION

In a free-market economy, prices play an important role. They help guide the so-called *invisible hands* (the term coined by the famous economist Adam Smith) in achieving an optimal allocation of resources. However, for the optimal allocation of resources to occur, the price should reflect the true or fundamental value of the asset. Otherwise, we will end up with the situation known as *market failure* where the resulting resource allocation is sub-optimal. Therefore, whether the price reflects the fundamental value of an asset is one of the fundamental questions in economics and finance. This question is of great interest to academicians, policymakers, business leaders as well as practitioners. In addition, it is important to understand the mechanism or process through which the fundamental value gets reflected in the price. The process is referred to as the *price discovery process.*

It is well recognized that the futures markets perform two central roles. Firstly, they provide instruments that can be used to hedge price risk. Secondly, they are supposed to be the markets where *price discovery* takes place.[1] Therefore, it is interesting to empirically analyze the price discovery contribution of the futures market in the presence of spot market. There are theoretical reasons to expect a significant contribution of the futures market in the price discovery process. For example, the futures price is expected to respond to new information faster than the spot price due to lower

[a] Monash University Malaysia, Selangor, Malaysia
*Professor of Banking and Finance at the School of Business, Monash University Malaysia, Jalan Lagoon Selatan, 47500 Bandar Sunway, Selangor Darul Ehsan, Malaysia. Email: keshab.shrestha@monash.edu We would like to thank Monash University Malaysia for the research support.

[1] U.S. law governing derivatives trading codifies these two purposes, stating that agricultural futures markets operate in the national public interest by "providing a means for managing and assuming price risks, discovering prices, or disseminating pricing information through trading in liquid, fair, and financially secure trading facilities," (Commodity Exchange Act 2012; Janzen and Adjemian (2017)).

transaction costs and ease of short selling associated with the futures contracts (Fleming et al. (1996), Silvapulle and Moosa (1999), and Shrestha (2014)). Furthermore, transactions can be quickly reversed in the futures markets. Therefore, we expect individuals with private or superior information (e.g., informed traders) to trade in the futures market to benefit from such information. Also, speculators prefer to hold futures contracts as they are not interested in the physical commodities per se and find it easy to offset futures positions. Finally, hedgers with storage constraints will also buy futures contracts instead of buying spot commodities. All these arguments lead us to expect that more price discovery should take place in the futures markets compared to the spot markets.

This brings us to the question of how the price discovery is measured. Garbade and Silber (1983) develop a model to measure the price discovery and introduce the concept of dominant and satellite markets where the dominant market is the place where the price discovery primarily takes place. Using a vector autoregressive (VAR) model, they also suggest a way of measuring the price discovery. They empirically analyze the price discovery for wheat, corn, oats, orange juice, copper, gold and silver, and find that the price discovery mainly takes place in the futures markets.

Since the spot and futures prices are normally found to follow unit-root processes with the two series being cointegrated, Schwarz and Szakmary (1994) extend the model suggested by Garbade and Silber (1983) by introducing the error correction term following the concept of cointegration developed by Engle and Granger (1987). Similar to Schwarz and Szakmary (1994), there are alternate methods of measuring price discovery which are also based on the concept of cointegration. One such price discovery measure, known as the component share (CS), is based on the Gonzalo-Granger permanent-temporary (PT/GG) decomposition proposed by Gonzalo and Granger (1995) (e.g., see Booth et al. (1999), Booth et al. (2002) and Harris et al. (2002)).

Hasbrouck (1995) proposes another measure, so-called *information share* (IS), where the price discovery or the information share of market *i* is based on the fraction of the long-run impact of the innovation represented by market *i*. However, this measure leads to the upper and lower bounds for the IS instead of a unique measure. Another limitation of IS is that it can only be applied to the case where the cointegrating relation between the futures and spot prices is one-to-one. Lien and Shrestha (2014) suggest a way to modify the Hasbrouck's IS that solve both the limitations. The modified IS, referred to as *generalized information share* (GIS), leads to a unique measure and can be used in cases where the cointegrating relationships among price series are not necessarily one-to-one.

There are existing studies that analyze price discovery related issued associated with agricultural commodities. For example, Walburger and Foster (1998) use more than 18 years worth of weekly data for 19 US feed cattle prices from 19 regional markets. Using a state-space formulation, they find that the spot prices do not move independently. They also find tight regional market price interrelationships with five and seven price discovery regions. Garcia et al. (2015) analyze the relationship between spot and futures prices for corn, wheat and soybean. They find that futures contract expired up to 35% above the cash price from 2005 to 2010. They theoretically explain the difference using the price of carrying physical grain and storage rate. Recently, Janzen and Adjemian (2017) perform price discovery analysis among four wheat futures contracts traded in Chicago, Kansas City, Minneapolis and Paris using high frequency data from 2008 to 2013. They use IS and CS to measure price discovery. They find that the price discovery mainly takes place in United States futures markets. However, the information share of the Paris market increased significantly after August 2010.

In this study, we empirically analyze the price discovery contributions of spot and futures markets for seven different agricultural related commodities. They include soybean, soybean meal, soybean oil, corn, wheat, cocoa and coffee.[2] We use daily data available in Datastream that ends on

[2] We also considered price discovery for cotton and hogs. However, these series were found to be stationary. Therefore, we could not perform the price discovery analysis because it requires the price series to be non-stationary.

29 December 2017.[3] We measure price discovery using both CS and GIS. We also use IS when we find the cointegrating relationship to be one-to-one. We find that the futures and spot prices of all seven commodities have single unit-root. We also find that the futures and spot prices are cointegrated with single cointegrating vectors. These two conditions are necessary to compute the GIS and CS measures. For four commodities, e.g., soybean, corn, cocoa and coffee, the cointegrating relationships are significantly different from one-to-one. However, for the remaining three commodities, the cointegrating relationships are found to be one-to-one. For these commodities, we are able to compute the Hasbrouck IS measure. We find that most of the price discovery takes place in the futures markets for all commodities with the exception of cocoa. For cocoa, the price discovery takes place both in the futures and spot markets. Our results show that futures markets play an important role in the price discovery process. Therefore, policy makers should encourage and facilitate the development of futures markets where such markets do not exit. We contribute to the existing literature in several ways. Firstly, we use the GIS measure to analyze the price discovery process that does not require the cointegrating relationship to be one-to-one. Secondly, we analyze seven different agricultural commodities to present comprehensive results. Thirdly, we use the most up to date data in our analysis.

The rest of the paper is organized as follows. In Section 2, we briefly describe the GIS and CS methods. We present the empirical results in Section 3. The paper concludes in Section 4.

INFORMATION SHARE MEASURES

In this section, we would briefly discuss the generalized information share (GIS) measure as well as CS measure. Both these measures are based on cointegrated unit-root processes. Let Y_t be an $n \times 1$ vector of unit-root series where it is assumed that there are $(n-1)$ cointegrating vectors which implies that the system consists of a single common stochastic trend (Stock and Watson (1988)). Therefore, the series have the following vector error-correction (VEC) representation (Engle and Granger (1987)):

$$\Delta Y_t = \prod Y_{t-1} + \sum_{i=1}^{k} A_i \Delta Y_{t-i} + \varepsilon_t, \qquad \prod = \alpha\beta^T \tag{1}$$

where β and α are $n \times (n-1)$ matrices of rank $(n-1)$. The columns of β consists of the $(n-1)$ cointegrating vectors and each column of α represents the adjustment coefficients. The matrix Π is decomposed in such a way that $\beta^T Y_t$ consists of $(n-1)$ vectors of stationary series. Let Ω denote the 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 vector moving average (VMA) representation (Hasbrouck (1995)):

$$\Delta Y_t = \Psi\,(L)\varepsilon_t \tag{2}$$

[3] We use daily data instead of high-frequency (intra-day) data because high-frequency spot prices are difficult to find for the commodities considered in the study. Some could argue that daily data constitutes too low frequency to perform price discovery. However, if one day is long enough for futures and spot prices to fully reflect the fundamental value, then the price discovery measures for these two prices should be approximately equal. However, as shown later, our empirical evidence shows that price discovery for the futures market ranges from approximately 71% to 100%, except for cocoa. Therefore, daily data seems to be useful at least for the agricultural commodities. However, for example, if we were to analyze the price discovery for foreign exchange markets, daily data may not be appropriate.

or, alternatively,

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

Due to the assumed nature of the cointegrating relationship among these unit-root series, the Engle-Granger representation theorem (Engle and Granger (1987)) implies the following (De Jong (2002) and Lehmann (2002)):

$$\beta^T\Psi\,(1) = 0 \; and \; \Psi\,(1)\alpha = 0. \tag{4}$$

Based on the above representations, Hasbrouck (1995) considers the $\Psi(1)\varepsilon_t$ to represent the long-run impact of innovations on the unit-root series. 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) Measure

Based on the above framework, we discuss the GIS measure.[4] Note that we have assumed n non-stationary series to have $(n-1)$ cointegrating vectors. Therefore, the cointegrating vectors represented by columns of matrix β can be written as follows:

$$\underset{(n-1)\times n}{\beta^T} = \left[\iota_{(n-1)} : -\Gamma_{(n-1)}\right] = \begin{bmatrix} 1 & -\gamma_1 & 0 & \dots & 0 \\ 1 & 0 & -\gamma_2 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \dots & -\gamma_{(n-1)} \end{bmatrix} \tag{5}$$

where $\Gamma_{(n-1)} = Diag\left(\gamma_1, \gamma_2, \dots, \gamma_{(n-1)}\right)$ and $\iota_{(n-1)}$ is an $(n-1)$ element column vector with all its elements equal to 1. For the above cointegrating structure to hold, the only requirement is that all the n unit-root series be driven by a single common stochastic trend. Let ψ_j^r be the j^{th} row of $\Psi(1)$. Then, above $(n-1)$ cointegrating relations imply the following (see equation (4)):

$$\psi_1^r = \gamma_{j-1}\psi_j^r, \qquad j = 2, \dots, n \tag{6}$$

Therefore, the *long-run impact* of innovations on the i^{th} series is given by

$$\psi_i^r\varepsilon_t = \psi_1^r\gamma_{i-1}^{-1}\varepsilon_t, \qquad i = 1, \dots, n \tag{7}$$

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

$$\psi_i^r\Omega\psi_i^{rT} = \sum_{j=1}^{n}\psi_{ij}^2\,\Omega_{jj} = \gamma_{i-1}^{-2}\sum_{j=1}^{n}\psi_{1j}^2\,\Omega_{jj} \tag{8}$$

[4]Please see Lien and Shrestha (2014) for detail.

The contribution of the innovation of series j to the total variance of the *long-run impact* of innovation on the i^{th} series is given by:

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

Note that, this measure is independent of i. Therefore, we can use the above measure as *Information Share* of series j for the case where the innovations are independent. For the general case, where the innovations are not independent, we can calculate the information share of the j^{th} series as follows:

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

where $\psi^G = \psi_1^r F^M,\ F^M = \left[G\Lambda^{-1/2} G^T V^{-1}\right]^{-1}$ and ψ_j^G is the j^{th} element of ψ^G. The information share measure given by equation (10) is referred to as generalized information share (GIS). It can be shown that the GIS is independent of ordering. Therefore, the GIS method leads to a unique measure of information share unlike the upper and lower bound for IS proposed by Hasbrouck (1995).[5]

Component Share (CS) Measure

Here, we briefly describe the CS method.[6] Since the empirical part of the paper deals with two unit-root series at a time, we will assume the number of series is two with one cointegrating vector, i.e., $n = 2$. In this case, the adjustment coefficient vector is denoted by $\alpha = (\alpha_1, \alpha_2)^T$.

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

$$f_t = \mu^T Y_t \tag{11}$$

where μ is an 2×1 permanent coefficient vector which can be shown to be orthogonal to the adjustment coefficient vector $\alpha, i.e., \mu = \alpha_\perp$. The normalized μ is given by

$$\mu = \alpha_\perp = (\mu_1, \mu_2)^T = \left(\frac{\alpha_2}{\alpha_2 - \alpha_1}, \frac{\alpha_1}{\alpha_1 - \alpha_2}\right)^T \tag{12}$$

[5] See Lien and Shrestha (2009) and Lien and Shrestha (2014) for detail on this issue.
[6] 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.

The component share of the first market (CS_1) and the component share of second market (CS_2) are given as follows

$$CS_1 = \mu_1 = \frac{\alpha_2}{\alpha_2 - \alpha_1} \;\&\; CS_2 = \mu_2 = \frac{-\alpha_1}{\alpha_2 - \alpha_1} \tag{13}$$

It is clear from equation (4) that the CS method uses information on Ψ (1), whereas the GIS uses information on both Ψ (1) and innovation covariance matrix Ω. Therefore, these two methods could lead to different conclusions. One practical limitation of the CS method is the fact that, in empirical studies, we may end up with the wrong sign for the estimates of the adjustment coefficients $\alpha_1 or\ \alpha_2$. For example, we expect to have α_1 to be negative and α_2 to be positive so that if $Y_{1,(t-1)}$ is above its long-run equilibrium value relative to $Y_{2,(t-1)}$, in the next period, $Y_{1,t}$ is supposed to decrease and $Y_{2,t}$ is supposed to increase so the system will move towards the long-run equilibrium where the long-run equilibrium is given by the cointegrating relationship. But, in empirical studies, we may end up with the estimates of $\alpha_1 or\ \alpha_2$ to have the wrong sign.[7] In this case, one of the CS measures ($CS_1\ and\ CS_2$) given by equation (13) will have negative value. In such cases, we use the convention by setting the parameter with the wrong sign to zero.

Finally, it is important to note that the VEC model represented by equation (1) can be estimated using the maximum likelihood method suggested by Johansen (1991). Once the model is estimated, we can use the likelihood ratio test to test the equality of the component shares for market 1 and market 2 as follows:

$$H_0 : \alpha_2 = -\alpha_1, \qquad H_1 : \alpha_2 \neq -\alpha_1$$

where the acceptance of the null hypothesis means that the difference in CS is not significant. Furthermore, we can also use the likelihood ratio test to see if the cointegrating vector is one-to-one by testing to see if the coefficient of the cointegrating vector γ_1 is – 1.0 as follows:[8]

$$H_0 : \gamma_1 = -1.0,\ H_1 : \gamma_1 \neq -1.0$$

[7] In this study, we find the estimates of α_1 to have wrong sign for four out of seven commodities analysed. However, the estimates of α_2 have the correct positive sign in all seven cases.

[8] Since the first element of the cointegrating vector associated with the first series is normalized to 1.0, the one-to-one cointegrating vector implies that the second element of the cointegrating vector associated with the second series $i.e., \gamma_1$ is equal to -1.0. The restriction can be tested using the likelihood ratio test.

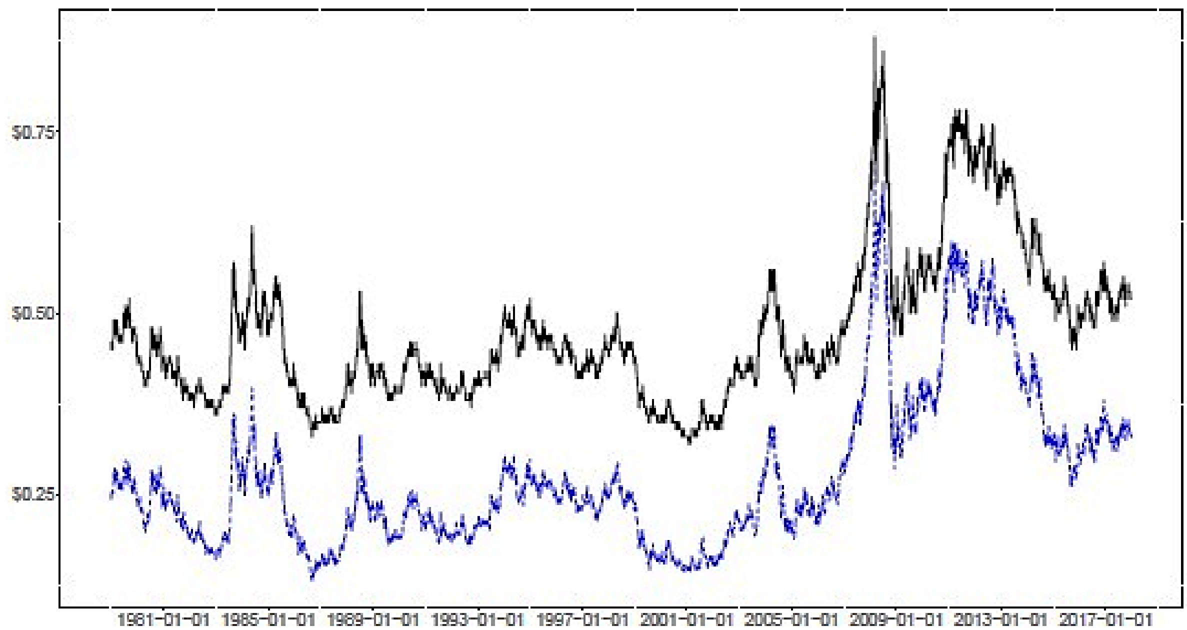

Figure 1: Daily soybean oil futures (dotted blue line) and spot price (solid line) in Dollars per pound from 2 January 1979 to 29 December 2017 where the spot price is shifted up by $0.20 by adding $0.20 to the spot price series.

EMPIRICAL RESULTS

In this study, we analyze the price discovery in the futures and spot markets for seven agricultural related commodities. They include markets for soybean, soybean meal, soybean oil, corn, wheat, cocoa and coffee. We use the daily data available in Datastream. Due to the availability of data in Datastream, the starting dates are different for different commodities. However, all of them have the same ending date, which is 29 December 2017. For the futures price, we use the nearest-month continuous series. The daily futures and spot prices for these commodities are plotted in Figures 1 through 7. As can be seen from the figures, there are large variations in the futures and spot prices allowing us to perform reliable statistical analysis. The information on the samples is given in Table 1. In all the analyses, we use the logarithm of the futures and spot prices.

In order to compute the generalized information share (GIS) and the component share (CS) measures, we need to establish the pre-conditions that the series under consideration are non-stationary, i.e., the series consists of single unit-roots. We use Phillips-Perron (PP) (Phillips and Perron (1988)) unit-root test on the level of the series and its first difference.[9]

[9] Augmented Dicky-Fuller tests lead to similar conclusions.

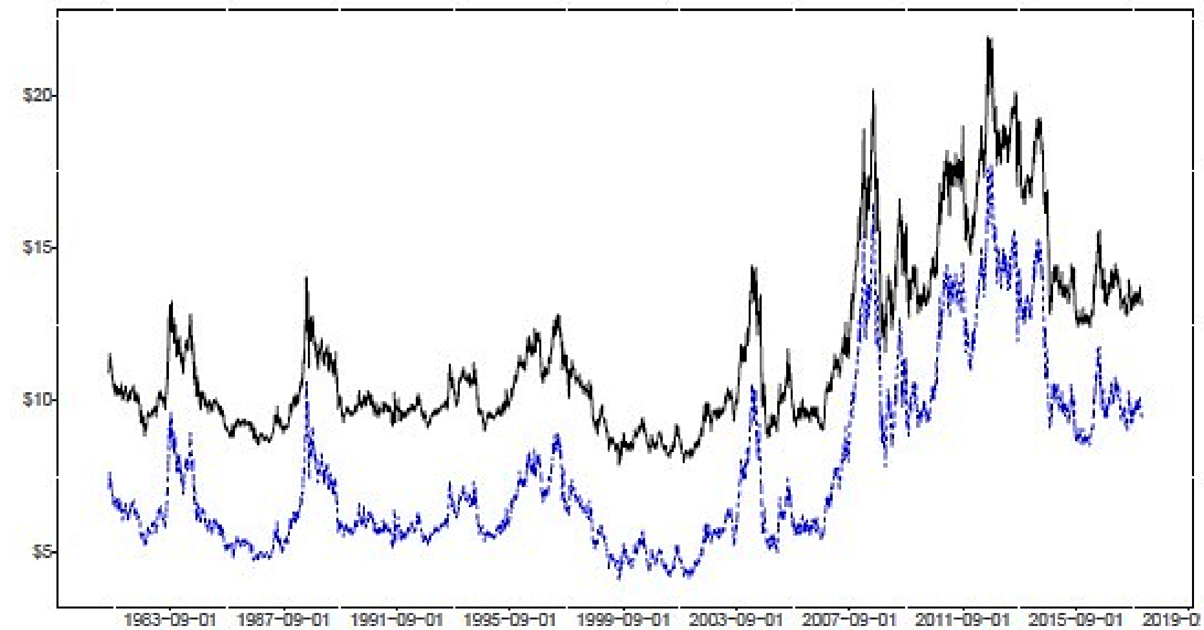

Figure 2: Daily soybean futures (dotted blue line) and spot price (solid line) in Dollars per bushel from 1st July 1981 to 29 December 2017 where the spot price is shifted up by $4.00 by adding $4.00 to the spot price series.

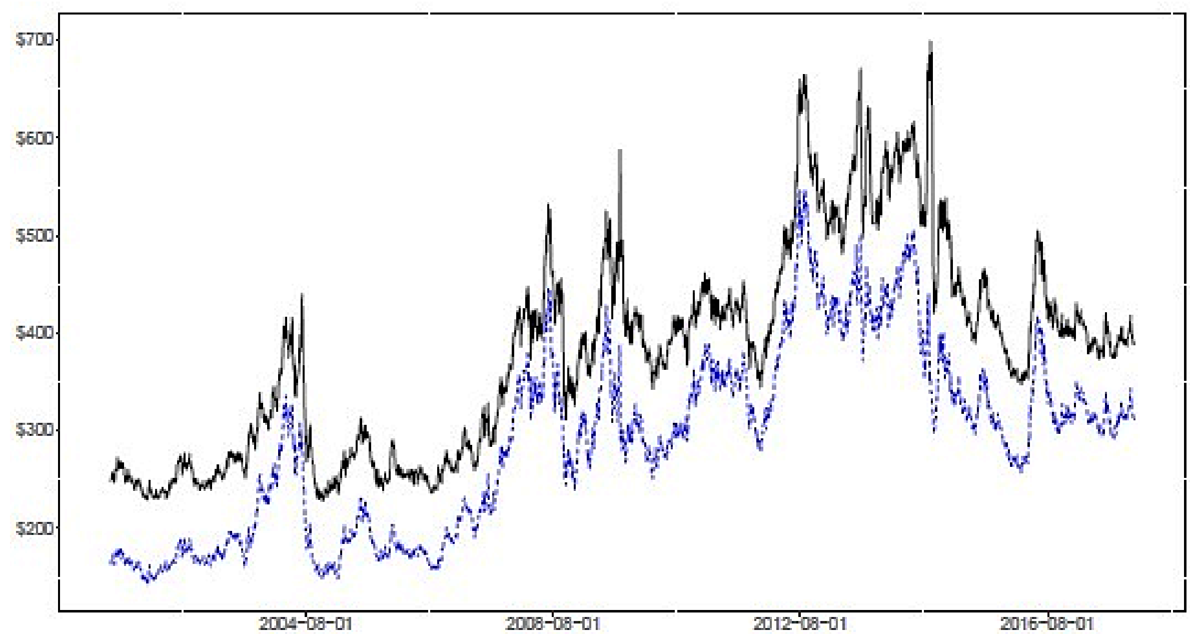

Figure 3: Daily soybean meal futures (dotted blue line) and spot price (solid line) in Dollars per metric ton from 1st June 2001 to 29 December 2017 where the spot price is shifted up by $80.00 by adding $80.00 to the spot price series.

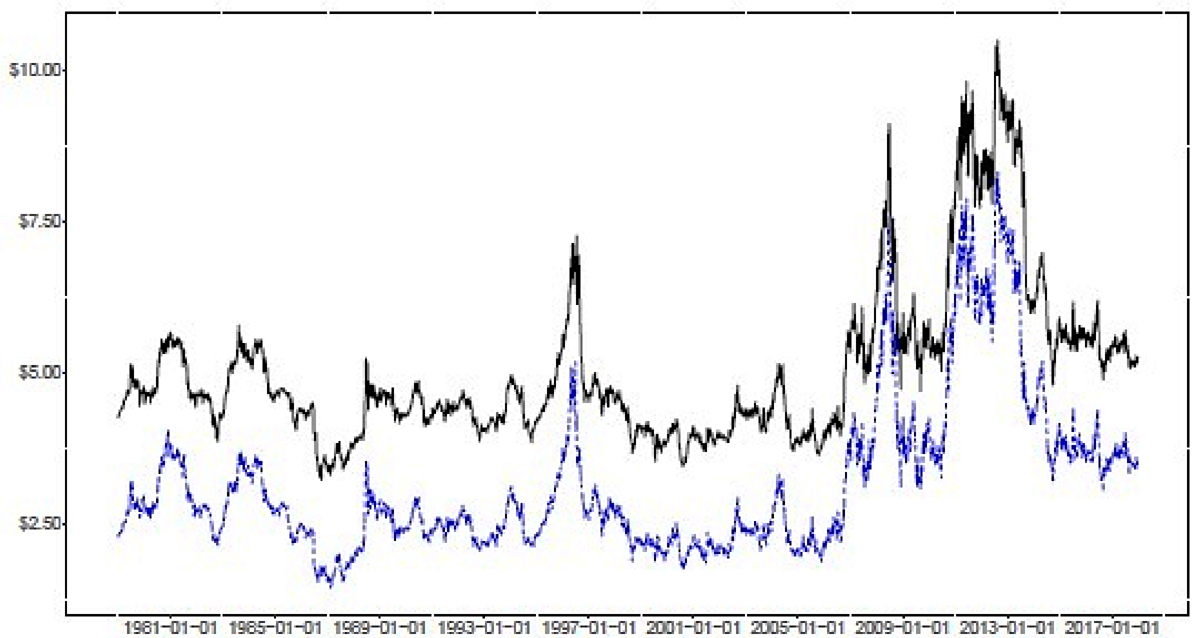

Figure 4: Daily corn futures (dotted blue line) and spot price (solid line) in Dollars per bushel from 2 January 1979 to 29 December 2017 where the spot price is shifted up by $2.00 by adding $2.00 to the spot price series.

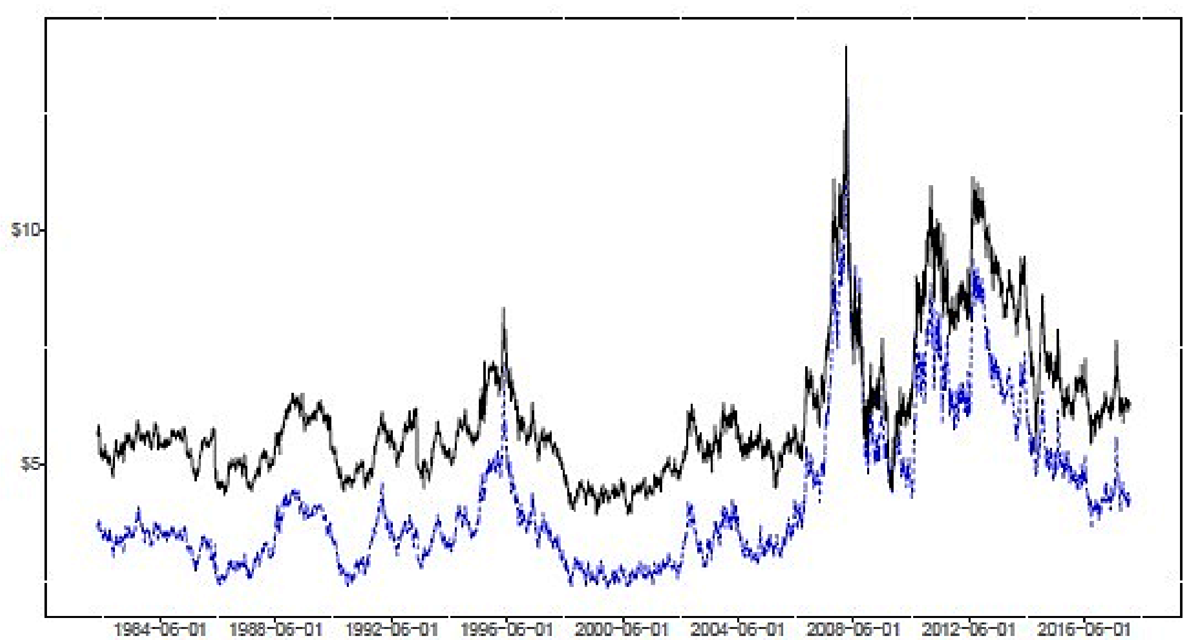

Figure 5: Daily wheat futures (dotted blue line) and spot price (solid line) in Dollars per bushel from 30 March 1982 to 29 December 2017 where the spot price is shifted up by $2.00 by adding $2.00 to the spot price series.

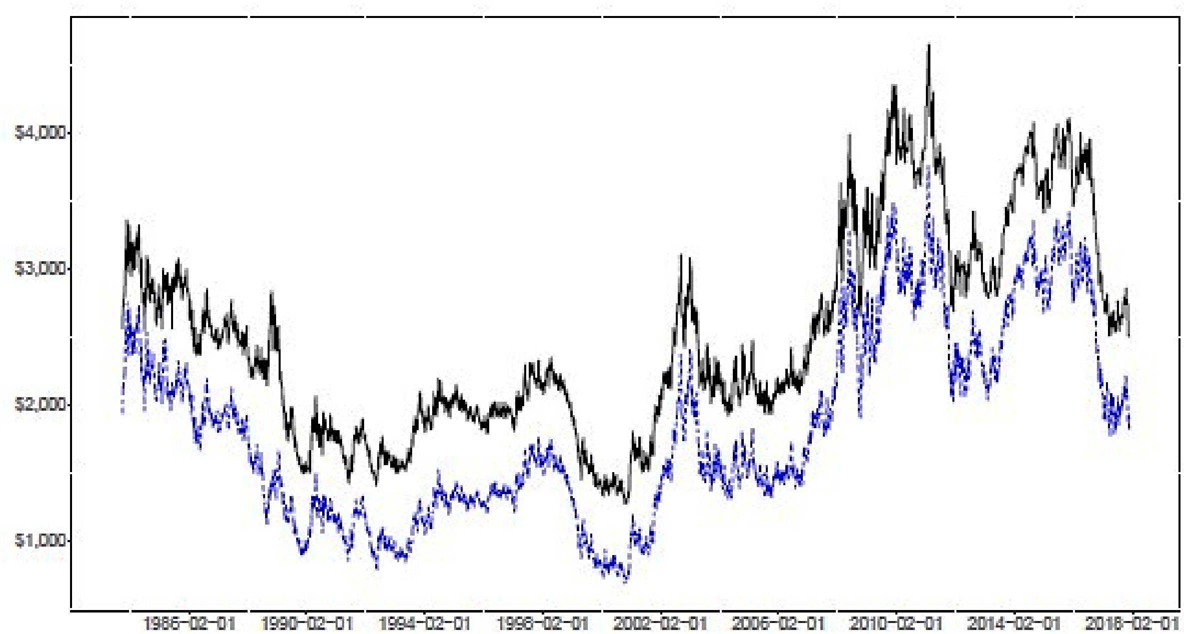

Figure 6: Daily cocoa futures (dotted blue line) and spot price (solid line) in Dollars per metric ton from 1st November 1983 to 29 December 2017 where the spot price is shifted up by $400.00 by adding $400.00 to the spot price series.

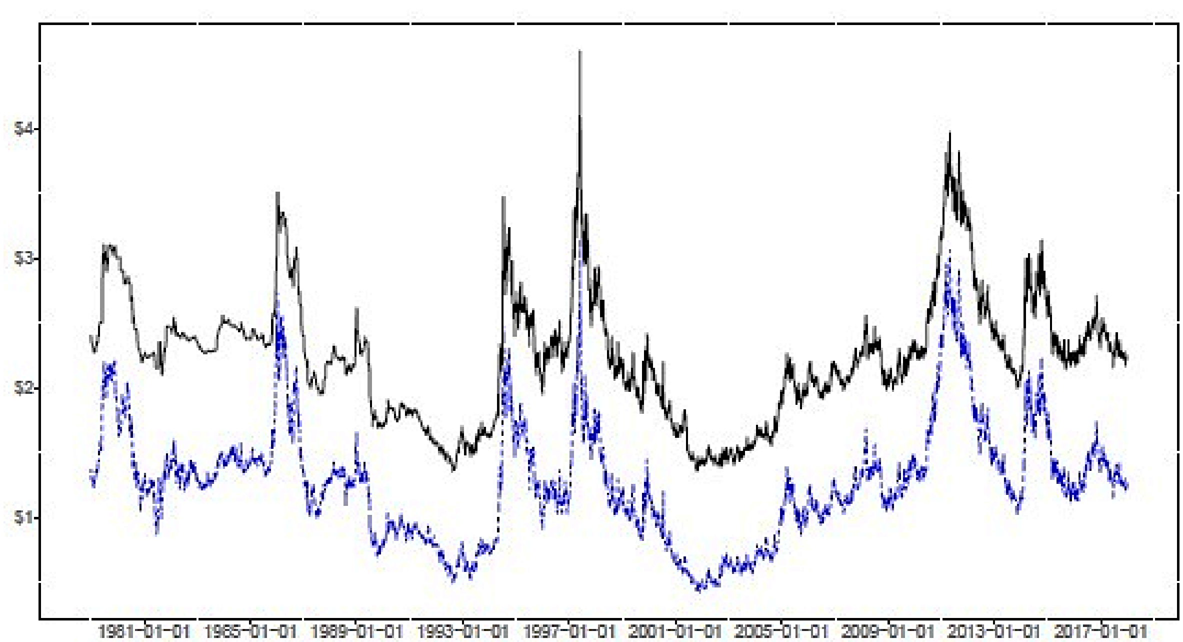

Figure 7: Daily coffee futures (dotted blue line) and spot price (solid line) in Dollars per bushel from 2nd January 1979 to 29 December 2017 where the spot price is shifted up by $1.00 by adding $1.00 to the spot price series.

Table 1: Commodities Used
This table shows the beginning date, end date and sample size for the seven futures contracts used in the study. The dates are in the YYYYMMDD format.

Commodity	Begin Date	End Date	Sample Size
1. Soybean oil	*19790102*	*20171229*	*9833*
2. Soybean	*19810701*	*20171229*	*9228*
3. Soybean meal	*20010601*	*20171229*	*4178*
4. Corn	*19790102*	*20171229*	*9848*
5. Wheat	*19820330*	*20171229*	*9017*
6. Cocoa	*19831101*	*20171229*	*8570*
7. Coffee	*19790102*	*20171229*	*9780*

The results are summarized in Table 2. All the unit-root tests for the level (i.e., logarithm of prices) are insignificant at the 5% level except for the spot price for wheat. This implies that all the series are non-stationary except the spot price for wheat. In order to test to see if the series have multiple unit-roots, we perform the PP tests on the first-differenced series. All the unit-root test statistics for the first-differenced series are highly significant. Therefore, we can conclude that all the series are non-stationary with single unit-root except for the spot price for wheat.

As the *uniformly most powerful* test for unit-root does not exist, it is important to use alternate tests to establish that the series considered are unit-root series. As unit-root is the null hypothesis under the Phillips-Perron, we also use the KPSS (Kwiatkowski et al. (1992)) test, which assumes stationarity as the null hypothesis. The results of the KPSS test are summarized Table 3. It is clear that the null hypothesis of stationarity is rejected for all series even at the 1% level of significance. Furthermore, the null hypothesis of stationarity for the differenced series cannot be rejected even at the 10% level of significance for all series. This is also true for the spot price for wheat. Therefore, we conclude that all the series considered have single unit-roots.[10]

[10] Even though, the PP test for the spot price for wheat is significant at the 5% level, the KPSS test for this series is significant even at the 1% level. Therefore, the PP test rejects the null hypothesis of unit-root at the 5% level and the KPSS test rejects the null-hypothesis of stationarity at the 1% level of significance for the spot price for wheat. Furthermore, the cointegration test, to be discussed later, indicates that there is one cointegrating relationship between the futures and spot prices for wheat. This provides an additional reason to conclude that spot price for wheat is a unit-root process. Otherwise, if the futures price is an unit-root process and the spot price is a stationary process, there should not exist any cointegrating relationship between these two series. Finally, unit-root is also consistent with the martingale principle implied by market efficiency (LeRoy (1989)).

Table 2: Phillips-Perron Unit-Root Test Results
This table summarizes the results of the Phillips-Perron (PP) Unit-Root Tests on the logarithm of spot and futures prices. The critical values are -2.57, -2.87, and -3.43 at the 10%, 5%, and 1% levels of significance respectively. ***, **, and * indicate the test statistic to be significant at the 1%, 5%, and 10% respectively.

	Log of Futures Price		Log of Spot Price	
	Level	**First Difference**	**Level**	**First Difference**
1. Soybean oil	-2.2446	-96.8683***	-2.4833	-127.9618***
2. Soybean	-2.1354	-94.5293***	-2.1310	-98.0426***
3. Soybean meal	-2.1363	-62.9435***	-2.1717	-60.6405***
4. Corn	-2.4632	-94.6503***	-2.5326	-98.7927***
5. Wheat	-2.3988	-93.2589***	-2.8991**	-100.9522***
6. Cocoa	-1.9955	-92.9231***	-1.9281	-103.9787***
7. Coffee	-2.8483*	-99.3973***	-2.4976	-98.4512***

Once it is established that each series has a single unit-root, next we want to perform tests to see if a cointegrating relationships between the futures price and the spot price exists for each of the seven pairs of futures and spot prices. We apply the Johansen (Johansen (1991)) cointegration test to test for the existence of the cointegrating relationship. The results are summarized in Table 4. We report both the λ_{max} and the trace statistics. Table 4 also reports the slope (i.e., coefficient, column (2)) and the intercept (column(3)) of the cointegrating vectors. In the cointegrating analysis, we use the logarithm of futures price as the first series and the logarithm of the spot price as the second series. Both the λ_{max} and trace statistics are highly significant for zero cointegrating vector even at the 1% level of significance.

Table 3: KPSS Unit-Root Test Results
This table summarizes the Kwiatkowski et al. (1992) (KPSS) unit-root tests on the logarithm of spot and futures prices. The critical values are 0.347, 0.463 and 0.739 at the 10%, 5%, and 1% levels of significance respectively. ***, **, and * indicate the test statistic to be significant at the 1%, 5%, and 10% respectively.

	Log of Futures Price		Log of Spot Price	
	Level	**First Difference**	**Level**	**First Difference**
1. Soybean oil	10.4578***	0.0421	9.3374***	0.0406
2. Soybean	11.5463***	0.0463	11.3570***	0.0458
3. Soybean meal	9.1342***	0.0594	9.1556***	0.0582
4. Corn	8.7435***	0.0336	7.6505***	0.0335
5. Wheat	10.3739***	0.0387	8.8763***	0.0271
6. Cocoa	9.5116***	0.0726	9.5412***	0.0743
7. Coffee	2.2197***	0.0345	2.0758***	0.0469

This is true for each of the seven pairs of the futures and the spot prices. Next, we test for the existence of at most one cointegrating vector. Since for this test, both the λ_{max} and trace statistics are identical, we only report the λ_{max} statistics. The., λ_{max} statistics, under `at most one' cointegrating vector, are insignificant even at the 10% level of significance for all seven commodities. Therefore, based on these statistics, we conclude that there exists a single cointegrating relationship between the futures and spot prices. This result is not surprising given that, theoretically, both the futures and spot prices are related to each other through the cost-of-carry arbitrage model. So far, we have established that the logarithm of spot and futures prices are non-stationary with single unit-root for all seven agricultural commodities. We have also established that there exists a single cointegrating vector between each pairs of the spot and futures prices. Therefore, we have satisfied the conditions necessary for the use of the GIS and the CS based information share measures in the analysis of the price discovery process. It is interesting to note that, for soybean, corn, cocoa and coffee, the cointegrating relationships are significantly different from one-to-one relationship needed to compute Hasbrouck IS.[11]

The computed GIS measures are reported in Table 5. Based on the GIS, it is clear that most of the price discovery takes place in the futures market for all commodities with the exception of cocoa. For soybean oil and soybean meal, close to 100% of the price discovery takes place in the futures markets. As for the cocoa, more price discovery seems to take place in the spot market. Next, we compute the CS based information share measure. As pointed out earlier, we expect the estimate of α_1 to be negative and the estimate of α_2 to be positive therefore any disequilibrium in the spot and futures prices on a day would be partially corrected on the following day through the appropriate change in the spot and futures prices. As reported in Table 4, all seven estimates of α_2 are positive as expected. However, the estimates of α_1 for soybean oil, soybean, soybean meal and corn are positive. Therefore, in the computation of the CS based information share, we replace the estimate of α_1 for these commodities with 0 in the computation of CS. The CS based information share measures are reported in Table 5. The CS based results are consistent with the GIS based results. We also perform statistical tests on the equality of CS for the futures and spot markets by testing the hypothesis $\alpha_2 = -\alpha_1$The results reported in Table 4 (columns (7) and (8), i.e., the last two columns) indicate the difference in the component shares between the futures and spot markets for cocoa and coffee are not significant even at the 10% level. However, the difference in component shares for all other commodities are significantly different at the 5% level.

As to the computation of the Hasbrouck IS, we find that the one-to-one cointegrating relationship for soybean oil, soybean meal and wheat cannot be rejected. For these commodities, the upper (high) and lower (low) bounds for the IS are reported in Table 5 (columns (6)-(9)). The Hasbrouck IS generally leads to the similar conclusions based on the GIS and CS measures. But, for the remaining four commodities, the one-to-one cointegrating relationship is rejected and, thus, the Hasbrouck IS cannot be computed.

[11] The one-to-one cointegrating relationship implies that the coefficient, γ_1 , is equal to -1.

Table 4: Cointegration Test Results

This table summarizes the results of Johansen Tests (e.g., λ_{max} and Trace tests) on the number of Cointegrating Vectors with lag length determined by the AIC criterion. The critical values are taken from Osterwald-Lenum (1992). ***, **, and * indicate the test statistic to be significant at 1%, 5%, and 10% level respectively. In the tests, we use the logarithm of the futures price to be the first series and the logarithm of the spot price to be the second series. The hypothesis that the cointegrating relationship is one-to-one, i.e., the coefficient, γ_1, is equal to -1, is tested. ***, **, and * next to the coefficient (column 2) indicate the test statistic to be significant at 1%, 5%, and 10% level respectively. The hypothesis that the CS of the futures and spot markets are equal, i.e., $\alpha_2 = -\alpha_1$, is also tested. ***, **, and * next to the α_1 and α_2 (columns 7 and 8) indicate the test statistic to be significant at 1%, 5%, and 10% level respectively.

Commodity	Coefficient γ_1	Intercept	Number of Cointegrating Vectors: None λ_{max}	None Trace	At most one λ_{max}	Adjustment Coefficients α_1	α_2
1. Soybean oil	-1.012	-0.031	25.222***	30.847***	5.625	0.0072***	0.0227***
2. Soybean	-0.978***	-0.065	128.367***	134.142***	5.775	0.0146***	0.0522***
3. Soybean meal	-0.955	-0.224	35.686***	40.635***	4.949	0.0136***	0.0280***
4. Corn	-0.943***	-0.107	90.448***	96.601***	6.153	0.0028***	0.0230***
5. Wheat	-1.049	0.027	33.902***	39.699***	5.797	-0.0002**	0.0095**
6. Cocoa	-1.066**	0.655	30.244***	33.935***	3.691	-0.0101	0.0081
7. Coffee	-0.850***	-0.751	69.005***	75.973***	6.968	-0.0073	0.0150

Table 5: Information Share Measure

This table summarizes the Generalized Information Share (GIS) for the futures and spot markets. It also reports the Component Share (CS) for the futures and spot markets. For commodities where the one-to-one cointegrating relationship cannot be rejected, the high and low Hasbrouck Information Shares (IS) are also reported.

	Generalized Information Share (GIS)		Component Share (CS)		Hasbrouck Information Share			
					Futures		Spot	
Commodity	Futures	Spot	Futures	Spot	High	Low	High	Low
1. Soybean oil	0.9977	0.0023	1	0	0.8921	0.8310	0.1690	0.1079
2. Soybean	0.8969	0.1031	1	0	NA	NA	NA	NA
3. Soybean meal	0.9975	0.0025	1	0	0.8198	0.7377	0.2623	0.1802
4. Corn	0.8433	0.1567	1	0	NA	NA	NA	NA
5. Wheat	0.8262	0.1738	0.9841	0.0159	0.9998	0.4382	0.5618	0.0002
6. Cocoa	0.4714	0.5286	0.4451	0.5549	NA	NA	NA	NA
7. Coffee	0.7066	0.2934	0.6711	0.3289	NA	NA	NA	NA

Based on these results, we conclude that price discovery mainly takes place in the futures markets except for cocoa. Even though, more price discovery seems to take place in the spot market for cocoa, still a significant level of price discovery also takes place in the futures market.[12] The evidence suggests that the futures markets have performed their price discovery role for the agricultural commodities considered in the study. Therefore, the policy makers should encourage the establishment of futures markets where such markets do not exist.

CONCLUSION

Prices play an important role in a free-market economy in guiding the allocation of resources to various uses. However, for the resulting allocation of resources to be optimal, the prices should reflect their fundamental values. Therefore, it is important to understand the price discovery process as well as the institutional arrangements that are supposed to improve the price informativeness. One of the objectives of the futures markets is to improve the price discovery process. In this study, we empirically analyze the contribution of futures markets to the price discovery process for seven agricultural commodities, which include soybean, soybean meal, soybean oil, corn, wheat, cocoa and coffee. We use daily data which include recent data up to 29 December 2017. We incorporate two different information share measures to analyze the price discovery process. The first one is the so-called generalized information share (GIS) proposed by Lien and Shrestha (2014) which is based on the original information share (IS) measure proposed by Hasbrouck (1995) and later on modified by Lien and Shrestha (2009). The Hasbrouck IS requires the cointegrating relationship between the futures and spot prices to be one-to-one. Whereas, the GIS does not impose this restriction. When we find the cointegrating relationship to be one-to-one, we also compute the upper and lower bounds for the Hasbrouck IS. The second measure, known as component share (CS), is based on the permanent-temporary decomposition proposed by Gonzalo and Granger (1995).

We find that the futures and spot prices of all seven commodities have single unit-roots. For all the commodities, we also find that the futures and spot prices are cointegrated with single cointegrating vectors. These two conditions are necessary to compute the GIS and CS measures. For four commodities, e.g., soybean, corn, cocoa and coffee, the cointegrating relationships are significantly different from one-to-one. However, for the remaining three commodities, the cointegrating relationships are found to be one-to-one. For these commodities, we are able to compute the upper and lower bounds for the Hasbrouck IS measure. We find that most of the price discovery takes place in the futures markets for all commodities with the exception of cocoa. For cocoa, more price discovery takes place in the spot market compared to the futures market. Our results show that the futures markets play an important role in price discovery process. Therefore, policy makers should encourage and facilitate the development of futures markets where such markets do not exit.

[12] It is important to note that, based on our empirical results, the daily frequency is useful for performing price discovery analysis for agricultural commodities. If the daily frequency is too low, both the futures and spot markets would have sufficient time to reflect the true or fundamental value of the commodity. In such cases, the information shares for both the futures and spot markets should be approximately equal. As our results indicate that this is not the case. Therefore, we can conclude that one day is sufficiently short period of time for performing the price discovery analysis at least for agricultural commodities. This may not be the case for other commodities like currencies.

REFERENCES

Baillie, R., Booth, G., Tse, Y., and Zabotina, T. (2002). Price discovery and common factor models. Journal of Financial Markets, 5:309-321.

Booth, G., Lin, J., Martikainen, T., and 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., and Tse, Y. (1999). Price discovery in german equity index derivatives markets. The Journal of Futures Markets, 19(6):619-643.

De Jong, F. (2002). Measures of contributions to price discovery: A comparison. Journal of Financial Markets, 5:323-327.

Engle, R. F. and Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. Econometrica, 55(2):251-276.

Figuerola-Ferretti, I. and Gonzalo, J. (2010). Modelling and measuring price discovery in commodity markets. Journal of Econometrics, 158:95-107.

Fleming, J., Ostdiek, B., and Whaley, R. E. (1996). Trading costs and the relative rates of price discovery in stock, futures, and option markets. Journal of Futures Markets, 16:353-387.

Garbade, K. and Silber, W. (1983). Price movements and price discovery in futures and cash markets. Review of Economics and Statistics, 65:289-297.

Garcia, P., Irwin, S. H., and Smith, A. (2015). Futures market failure? American Journal of Agricultural Economics, 97:40-64.

Gonzalo, J. and Granger, C. (1995). Estimation of common long-memory components in cointegrated systems. Journal of Business and Economic Statistics, 13:27-35.

Harris, F., McInish, T., and Wood, R. (2002). Security price adjustment across exchanges: an investigation of common factor components for dow stocks. Journal of Financial Markets, 5:277-308.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. Journal of Finance, 50:1175-1199.

Janzen, J. P. and Adjemian, M. K. (2017). Estimating the location of world wheat price discovery. American Journal of Agricultural Economics, 99:1188-1207.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. Econometrica, 59:1551-1580.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., and 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.

Lehmann, B. (2002). Some desiderata for the measurement of price discovery across markets. Journal of Financial Markets, 5:259-276.

LeRoy, S. F. (1989). Efficient capital markets and martingales. Journal of Economic Literature, 27:1583-1621.

Lien, D. and Shrestha, K. (2009). New information share measure. The Journal of Futures Markets, 29(4):377-395.

Lien, D. and Shrestha, K. (2014). Price discovery in interrelated markets. The Journal of Futures Markets, 34(3):203-219.

Phillips, P. and Perron, P. (1988). Testing for a unit root in time series regression. Biometrica, 75:335-346.

Schwarz, T. and Szakmary, A. (1994). Price discovery in petroleum markets: Arbitrage, cointegration, and the time interval of analysis. Journal of Futures Markets, 14:147-167.

Shrestha, K. (2014). Price discovery in energy markets. Energy Economics, 45:229-233.

Silvapulle, P. and 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.
Stock, J. H. and Watson, M. W. (1988). Testing for common trends. Journal of the American Statistical Association, 83:1097-1107.
Walburger, A. M. and Foster, K. A. (1998). Determination of focal pricing regions for u.s. fed cattle. American Journal of Agricultural Economics, 80:84-95.