Price Discovery in Agricultural Futures Markets: Should We Look Beyond the Best Bid-Ask Spread?

Mehdi Arzandeh and Julieta Frank*

Paper presented at the CAES annual meeting, Graduate Students Competition
Montreal, QC. June 18-21, 2017

---

*Mehdi Arzandeh is a PhD candidate in the Department of Economics (umarzand@myumanitoba.ca) and Julieta Frank is an Associate Professor in the Department of Agribusiness and Agricultural Economics (Julieta.Frank@umanitoba.ca) at the University of Manitoba.

This research was supported by the Social Sciences and Humanities Research Council of Canada. This research was enabled in part by support provided by WestGrid (www.westgrid.ca) and Compute Canada Calcul Canada (www.computecanada.ca).

Price Discovery in Agricultural Futures Markets: Should We Look Beyond the Best Bid-Ask Spread?

Price discovery is defined as the incorporation of information to prices through the actions of traders. Previous studies in financial markets have found evidence that informed traders may submit limit orders instead of market orders. If so, the steps of limit order book (LOB) beyond the best bid and best ask spread (BAS) contain valuable information and contribute to price discovery of the underlying asset. This is the first attempt to examine the informativeness of the LOB beyond the BAS for agricultural commodities. We reconstruct the LOB using market depth data and use three information share approaches to test whether the steps of LOB beyond the BAS contribute to price discovery in agricultural commodity markets. This is done for five major agricultural commodities namely live cattle, lean hogs, corn, wheat, and soybeans as well as the CME E-mini S&P 500. We find that a substantial market depth exists at the steps beyond the best bid and ask prices in the futures markets. The results of the three information share measures show that the steps of the LOB beyond the BAS contribute by over 27% to price discovery of futures contracts. Across agricultural commodities, the steps of the LOB beyond the BAS have more information for grains than meats. Moreover, we find that the steps closer to the top of the book, relative to the steps farther, contain more information. These findings suggest that informed traders in futures electronic markets actively use limit orders with price steps beyond the BAS and especially the steps near the top of the book. The results also show that for E-mini S&P 500, the steps closer to the top of the book contain more information at the beginning and the end of the week whereas steps farther have more information in the middle of the week.

Keywords: Futures Markets, Information Share, Commodity Markets, Electronic Trading, Limit Order Book

Introduction

Agricultural commodity futures were traditionally traded in the open outcry pit, however, over the past decade there has been a major shift to trading on the electronic platform. Grain and livestock futures contracts trading electronically weighed less than five percent of overall trade in 2006 and grew to over eighty and ninety percent, respectively, in 2011 (Irwin and Sanders 2012). Today the Chicago Mercantile Exchange (CME) Group, the largest futures contracts open interest exchange, has migrated its agricultural futures trading to the electronic platform. The electronic system differs significantly from the traditional open outcry system. One major difference is the presence of the limit order book (LOB) in the electronic system, which contains actual bid and ask prices and their corresponding volumes at different steps (Gould et al., 2013).

Trades in the electronic platform are conducted through a computerized system where all traders submit their orders with the number of contracts they want to trade and their intended prices. Traders can buy or sell contracts at existing market prices. If the price for which a trader intends to sell (buy) a contract is less than or equal (greater than or equal) to the price for which another trader intends to buy (sell) the contract, the trade will take place. This is also known as a market order. If, however, a trader's bid price is lower than the lowest ask price for the contract (i.e., the best ask), the order will remain active in the exchange electronic system on the bid side until it is matched or cancelled (or expired if it is a futures contract). The bid side, thus, can be thought of as the demand side for the underlying contract. Similarly, if a trader's ask price is higher than the highest bid price (i.e., the best bid), it remains active on the ask side until it is matched or cancelled (or expired). The ask side can be considered the supply side for the contract. The orders resting in the system are called limit orders and the system storing these orders is the LOB. At any point in time, the LOB contains all the resting orders on the demand and supply sides at different price steps. In the LOB, the best bid and best ask are the highest bid and the lowest ask prices, respectively, at that point in time which are referred to as "the top of the book". The difference between the lowest ask and the highest bid is called "the spread" or bid-ask spread (BAS). The other bids and asks are resting in descending and ascending order beyond the best bid and best ask, respectively, in the LOB.

The information contained in the LOB has been the subject of much controversy. If informed traders use limit orders, their information is presumably reflected in the book. If, however, informed traders use market orders, the orders in the book may not contain any of their private information. Several studies on the type of orders used by informed traders and the extent to which prices in the LOB carry information about the efficient price have been conducted, however the results are mixed (some examples are Harris and Panchapagesan 2005, Kaniel and Liu 2006, and Madhavan et al. 2005). In addition, only few of those studies analyze futures markets and none of them examine agricultural commodities.

One of the most important functions of futures markets is price discovery, which is the process of incorporating market participants' new information into market prices. Many studies have examined the contribution of related price series, such as securities trading in different markets or spot and futures prices of a commodity, to an underlying common efficient price. Hasbrouck's (1995) Information Share (IS) and Gonzalo and Granger's (1995) Permanent-Transitory (PT) measures have been widely used to assess the contribution of the related series to price discovery (some examples are De Jong 2002, Huang 2002, Booth et al. 2002, Chu et al. 1999, and Harris et al. 2002). Cao et al. (2009) study the information share of the steps of the LOB beyond the BAS on the underlying price. By looking at 100 active Australian stocks, they find that a share of about 22 percent of the price discovery can be attributed to the steps of the LOB beyond the best bid and best ask, whereas the remaining 78 percent is contributed by the best bid and ask and the last transaction price. However, the informational content of the LOB in agricultural markets may differ considerably for a variety of reasons. Markets for futures contracts are different from spot markets because many market participants trade in futures markets for the purpose of hedging and risk management. This implies that trading algorithms which are practiced in the two markets can be different. Agricultural commodity trading in futures markets can be different from trading other contracts due to differences in market characteristics such as tick size, availability of the commodity, etc. Even though much research on price discovery has been done for agricultural futures markets in the traditional outcry system, none has been done for the electronic market at the microstructure level.

The objective of the research is to assess the informational content of the LOB beyond the BAS in agricultural futures markets. We reconstruct the full LOB and compute both the BAS at the best quotes and the bid and ask at subsequent steps of the LOB beyond the best quotes. The informational content of the order book is then assessed by estimating the contribution of each of these series to price discovery. Cao et al. (2009) examines the contribution of the LOB to price discovery for the stock markets. This study focuses on electronic futures markets. The study, specifically, is performed using nearby contracts for five major agricultural commodities, namely live cattle, lean hogs, corn, wheat, and soybeans, as well as the popular E-mini S&P 500 from the CME Group. The products under study cover majorly traded agricultural commodities and, in order to compare the results with other actively traded futures contracts for which more research exists, the E-mini S&P 500 is also examined. Agricultural commodities are generally less traded and their market characteristics may be different from those of other products. Grain traders have access to nine steps beyond the best BAS and livestock traders have access to four steps beyond the best BAS in real time. Therefore, a better understanding of the contribution of the LOB to price movements may play a fundamental role in developing their trading algorithms and strategies.

One difficulty in assessing the information content of the LOB is that both Hasbrouck's IS and Gonzalo and Granger's PT measures use different approaches to estimating the contribution of price series to the common price, and there is no consensus in the literature favoring either estimate. While the PT measure is unique, it ignores the correlations between different price series (Hasbrouck 2000; Ballie et al. 2002). On the other hand, while IS accounts for this correlation, it is not unique as it is sensitive to the ordering of price series in the model. Lien and Shrestha (2009) propose an alternative measure, the modified information share (MIS), which uses an eigenvector factorization of the correlation matrix of residuals and thus is independent of the ordering. Here we estimate IS, PT, and MIS to assess the information content of the LOB.

Background

Information Contained in the LOB

The evidence on the extent to which price steps beyond the BAS carry information about the efficient price is mixed. Glosten (1994), Rock (1996), and Seppi (1997) argue that informed traders favor and actively submit market orders, suggesting that the LOB beyond the best bid and offer contains little information. However, Bloomfield et al. (2005) use an experimental market setting and find that in an electronic market, informed traders submit more limit orders than market orders. This suggests that key trader information is contained in the book. In the context of stock trading at the New York Stock Exchange (NYSE), Harris and Panchapagesan (2005) show that the imbalances in the limit buy and sell orders in the book have information regarding the short run price movements and that NYSE specialists benefit from it by buying for their own account when the book is heavy on the buy side and sell when it is heavy on the sell side, especially for more active stocks. Kaniel and Liu (2006) show that informed traders prefer limit orders, and that limit orders convey more information than market orders. Baruch (2005) provides a theoretical model showing that an LOB improves liquidity and information efficiency of prices. Boehmer et al. (2005) find that the deviations of transaction prices from the efficient prices became smaller after the NYSE's adoption of the LOB system. In contrast, Madhavan et al. (2005) find larger spreads and higher volatility after the Toronto Stock Exchange disseminated the top four price steps of the limit-order book in April 1990. More recently, Biais et al. (2015) and Martinez and Roşu (2011) develop theoretical models to compare algorithmic traders and humans in the informativeness of prices. Both studies suggest that algorithmic traders are advantageous compared to humans due to their quicker response to new information and that algorithmic traders use market orders to exploit their information. Hautsch and Huang (2012), on the other hand, estimate impulse response functions for thirty stocks traded at Euronext Amsterdam and find that limit orders, especially for orders posted on up to two steps beyond the market price, have a significant effect on quote adjustments. Moreover, Eisler et al. (2012) find further support on the effect of limit orders on market prices and Cont et al. (2014) argue that the order flow imbalances between supply and demand at the best bid and ask spread are the main driving force behind market price changes.

Research on agricultural markets at the microstructure level has relied on the transaction prices to study the best bid and ask. Examples of the earlier studies are Brorsen (1989), Bryant and Haigh (2004), and Hasbrouck (2004). Among the more recent studies, Frank and Garcia (2011), Shah and Brorsen (2011), and Martinez et al. (2011) used trade data to estimate BAS in measuring the cost of liquidity and comparing open outcry to electronic trading for different agricultural commodity markets. Wang et al. (2014) reconstructed the best bid and ask steps of a limit order book to study the liquidity costs in corn futures markets. Using CME Group RLC market depth data[1], Aidov (2013) and Aidov and Daigler (2015) reconstruct the five-step LOB for futures contracts of the 10-Year U.S. Treasury note, corn, light sweet crude oil (WTI), euro/U.S. dollar, yen/U.S. dollar, and gold futures to study the characteristics of market depth in electronic futures market such as duration, symmetry, and equality of depth. They find that the steps beyond the BAS contain a large amount of depth for all the futures contracts studied. Aidov (2013) derives the market depth from the five bid and ask steps to, firstly, study the relationship between the market depth and the bid-ask spread and secondly, to examine the link between the transitory volatility and the market depth. His results indicate a negative relationship between the five-step market depth and the spread. His results also suggest a decrease in market depth following an increase in volatility. He concludes that market participants in the U.S. electronic futures market actively manage depth along the LOB. Our study extends the previous literature on agricultural commodity electronic trading in some important ways. First, we reconstruct the LOB for four other major agricultural commodities i.e. live cattle, lean hogs, wheat, and soybeans besides corn and for the E-mini S&P 500 futures contracts. Second, this is the first study that examines the contribution of the LOB beyond the BAS to price discovery in agricultural commodity markets.

Information Share Measures

Hasbrouck's IS and Gonzalo and Granger's PT are the two most well-known information share measures used in the majority of the literature (some examples are Anand and Subrahmanyam 2008, Chen and Gau 2010, Frijns et al. 2010, Korczak and Phylaktis 2010, Anand et al. 2011, Fricke and Menkhoff 2011, Liu and An 2011, Chen and Chung 2012, Chen and Choi 2012, Rittler 2012, and Chen et al. 2013). However, some weaknesses have been identified in both measures and therefore efforts have been made to generate new measures. The IS measure is problematic because of its non-uniqueness and sensitivity to price ordering in the estimation. To overcome this problem, Hasbrouck (1995) has proposed calculating upper and lower bounds for

[1] RLC market depth data was discontinued in 2009 and the current format of the CME Group market depth data is FIX which is the format used in the present study. The period of study in Aidov (2013) and Aidov and Daigler (2015) for different contracts ranges from January 2008 to March, April, or October 2009, depending on when the RLC data was discontinued for the specific contract.

the information shares of price series. High frequency data minimizes the correlation between price series and results in close lower and upper bounds. However, the bound widens as the contemporaneous correlation of disturbances across the price series increases, making the discrepancies between the orderings become large enough to deemed inference on the information share of prices unreasonable (Tse, 1999; Huang, 2000; Harris et al. 2002). Lien and Shrestha (2009) avoid the order-dependency problem by using an eigenvector factorization of the correlation matrix of residuals (instead of a Cholesky factorization) in their modified IS (MIS) new measure. They later extend their MIS measure for the cases where the price discovery contribution of different but related financial securities are analyzed such as price discovery in markets for different securities issued by the same firm and propose the Generalized Information Share, GIS (Lien and Shrestha 2014). Yan and Zivot (2010) and Putnins (2013) showed that the information shares calculated using IS and PT do not account for the different levels of noise in the price series. They argue that this may result in misleading measures of information share and develop a new information share metric, the Informational Leadership (IL) by combining IS and PT. The IL is, however, applicable to a two price series setting. Moreover, different level of noise arises when studying, for example, price discovery of an asset traded in different markets with different characteristics such as minimum tick size, inventory management, or other market imperfections and microstructure frictions. Another measure of price discovery was developed by Grammig and Peter (2013) to address the IS problem of non-uniqueness, particularly for longer sampling intervals. They assume a multivariate mixture distribution to develop the tail-dependent information shares (TLS). Like MIS, TLS follows from Hasbrouck (1995) contribution of a price series variance to the variance of the efficient price as the measure of the series information share by means of reduced VECM long run impact coefficients. However, the variance decomposition under TLS is performed using a VECM which is extended by the mixture parameters and estimated by a two-step process. This, unlike IS, results in an order neutral measure and is claimed to be superior to IS and PT when correlations of price innovations in the tails differ from those in the center of the distributions. Lien and Wang (2016) compare the IS upper and lower bound midpoint with the two unique, more recent, information shares of MIS and TLS. They find that TLS performs poorly for the simulated data even when the underlying assumptions of the approach are met. Moreover, their results show that MIS at most marginally improves the information share computed by the IS midpoint. They, therefore, support the use of the IS midpoint as a method of computing the information shares of different price series.

Data

Date and Time

We estimate the information contained in the limit order book for live cattle, lean hogs, corn, wheat, soybeans, and E-mini S&P 500 nearby futures contracts trading in the CME Group for the period of November 23, 2015 to March 31, 2016. The total number of the LOB updates for the nearby futures contracts during this period is 19,280,306 for live cattle, 14,487,393 for lean hogs, 48,400,937 for corn, 38,690,537 for wheat, 106,390,881 for soybeans, and 575,528,486 for E-mini S&P 500. The LOB is updated when a trader submits a market order, a limit order or a deletion/cancellation order. Corn, wheat, and soybeans futures contracts trade in two sessions, 8:30 am to 1:20 pm CT (morning session) and 7:00 pm to 7:45 am CT (evening session). We use data from the morning session from Monday to Friday only, due to the low volume traded in the evening session and on Sunday. Live cattle and lean hogs futures contracts trade in one session only, 8:30 am to 1:05 pm CT. For the E-mini S&P 500 futures contract we use the most active daily trading hours, from 7:30 am to 3:15 pm CT, from Monday to Friday.[2] There is no trading on the CME Group on Saturdays and the two federal holidays of Jan. 18 and Feb. 15. We also remove the data for Sundays and the days for which there are extended trading halts. The latter is mostly the case for a few futures contracts with partial pre-holiday (a day prior) and post-holiday (a day after) trading with extended trading breaks. After thinning the data and restricting the data to the nearby futures contracts, we are left with total LOB updates of 6,961,908 for live cattle, 6,176,794 for lean hogs, 25,305,597 for corn, 21,285,007 for wheat, 42,625,431 for soybeans, and 386,421,232 for E-mini S&P 500.

Roll Dates

Traders in the CME Group futures can choose to roll their futures positions from one futures contract month to the next at any time. They roll forward their futures positions before the futures contracts are very close to termination and becoming very illiquid. Traditionally, traders in the CME Group roll forward expiring futures contracts eight calendar days before the contract expiry (i.e., the "roll date"). The eight calendar day roll period seems to be a good approximation of when traders roll their futures position to the next contract for the E-mini S&P 500. However, the period proves to be too short for the agricultural commodity futures contracts. Traders of the agricultural commodities start to roll their position considerably earlier than eight calendar days prior to expiration. We use the following rule to find the roll dates for each agricultural commodity and each contract month. We define a roll date for the current contract as the date when its aggregate volume traded for two consecutive days falls below that of the second nearest contract.

In figures 1 to 6, the daily aggregate volume for the near and active contracts are plotted for the six futures markets under study. The vertical lines show the roll dates for each contract, that is, when the highest aggregate volume traded switches from one contract to the next. It can be seen in the figures that the roll period for agricultural commodities is considerably longer than that of the E-mini S&P 500.[3]

[2] These hours correspond to 8:30 am – 4:15 pm ET. In GMT, this period is from 1:30 pm to 9:15 pm for before 13 Mar. 2016 (start of the daylight saving) and it is from 12:30 pm to 8:15 pm on and after 13 Mar. 2016. There is also a 15-minute trading halt from Monday to Friday at 3:15 pm - 3:30 pm CT.

[3] We also considered alternative rules where we examined the daily number of trade price changes and the daily average duration of price changes to determine the roll dates. These rules almost always result in the same roll dates as the case of aggregate volume rule. In the isolated exceptions, the roll date is one day before the roll date defined by the aggregate volume rule.

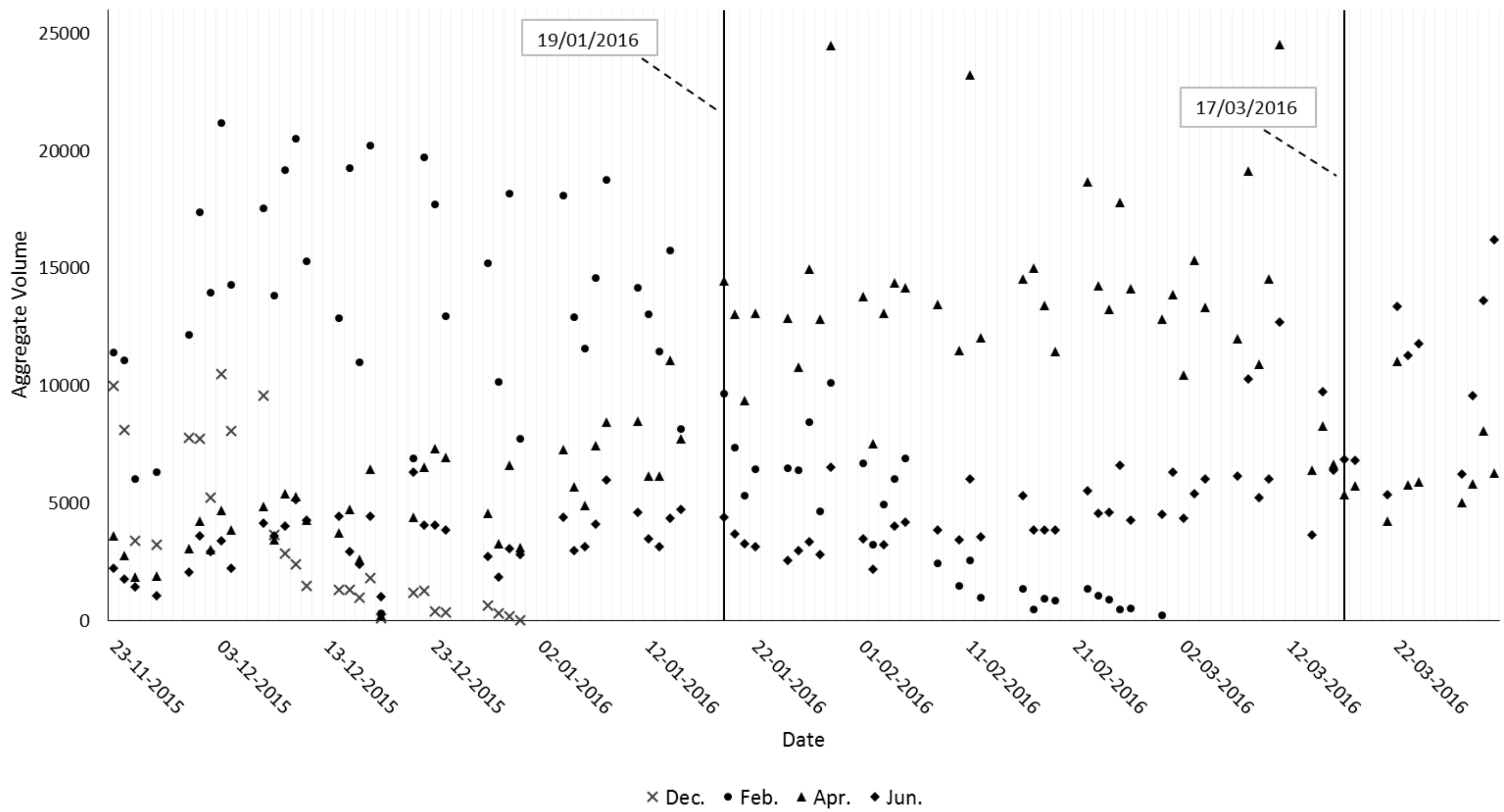

Figure 1: Roll dates for live cattle

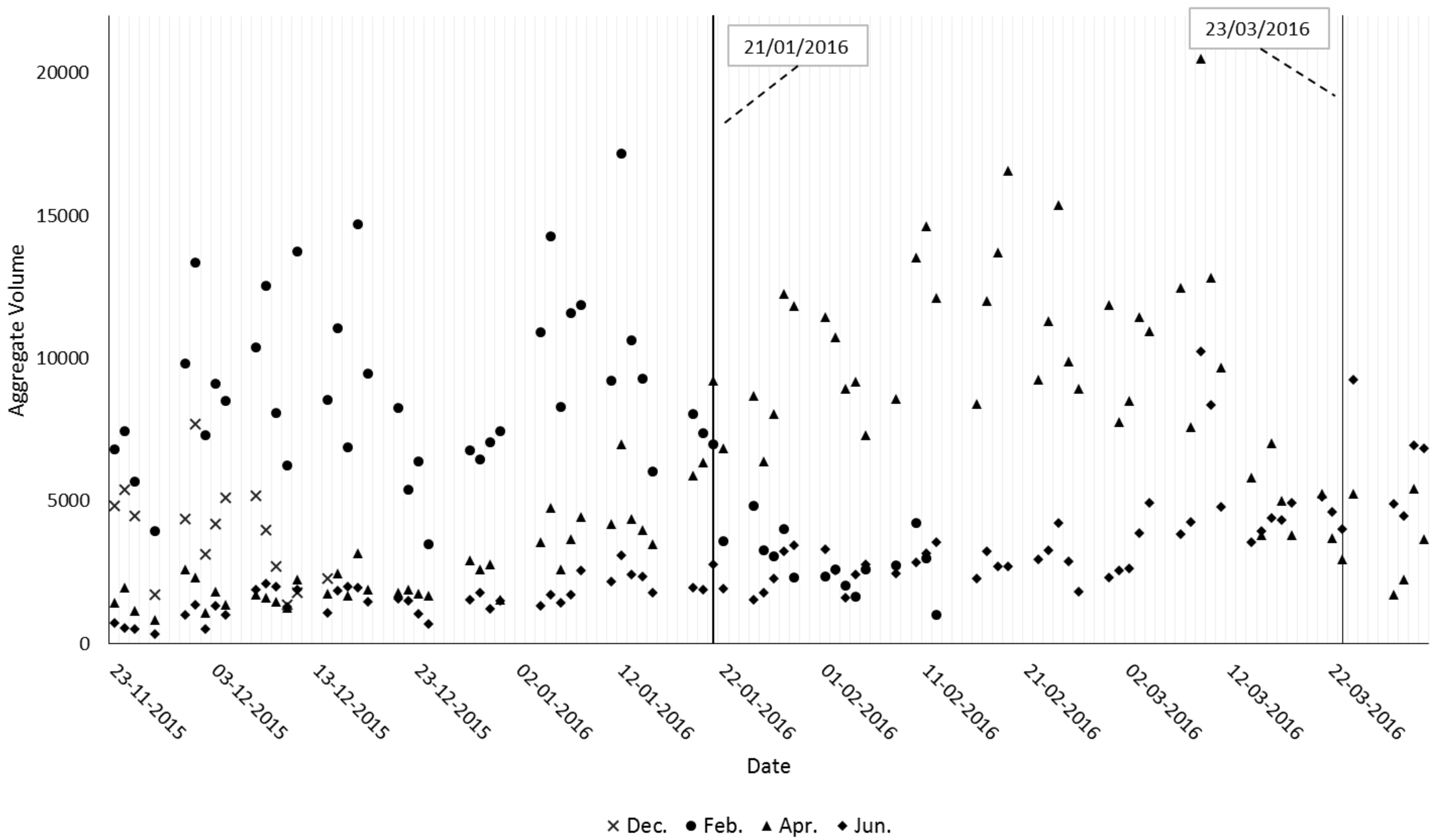

Figure 2: Roll dates for lean hogs

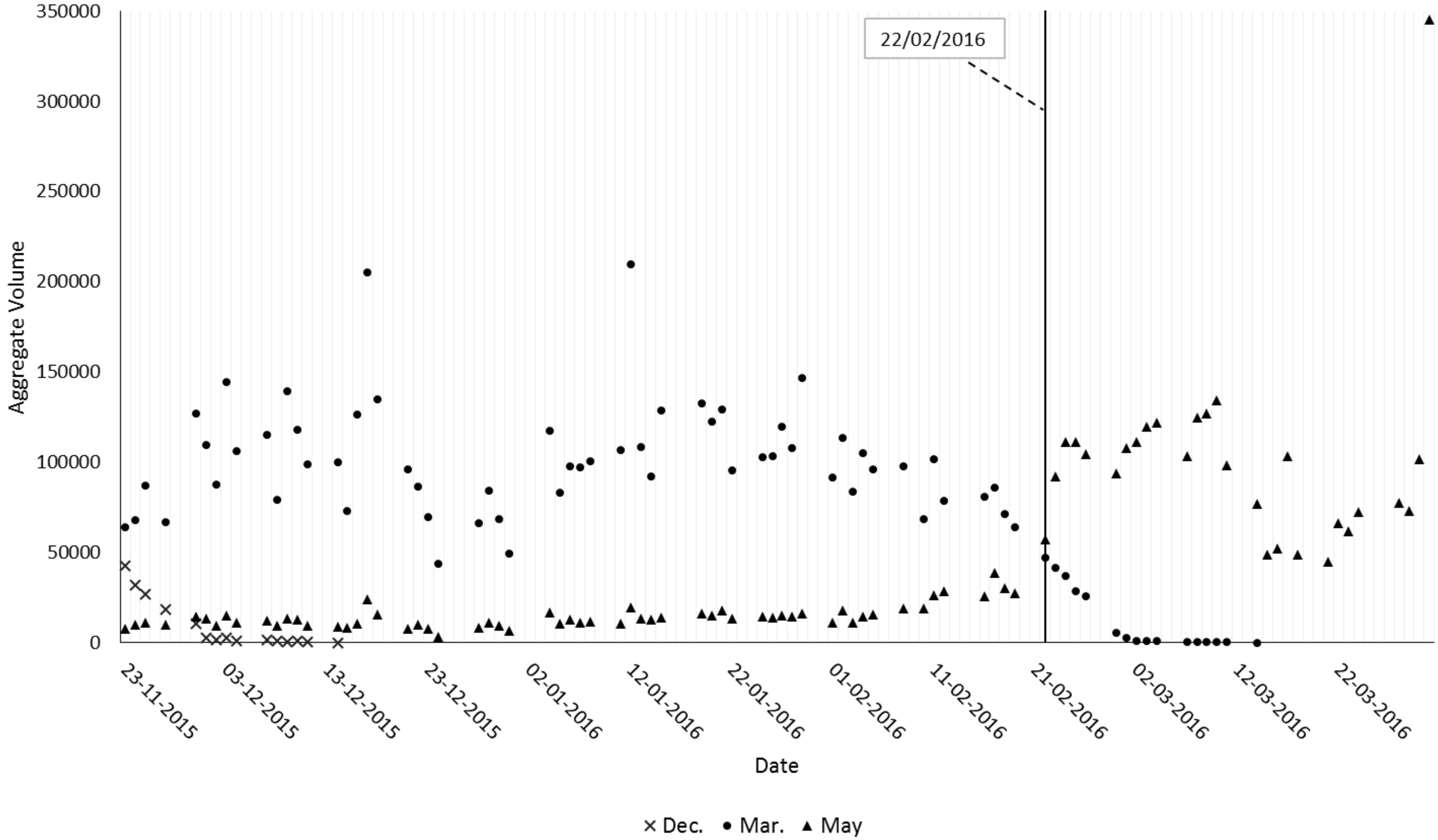

Figure 3: Roll dates for corn

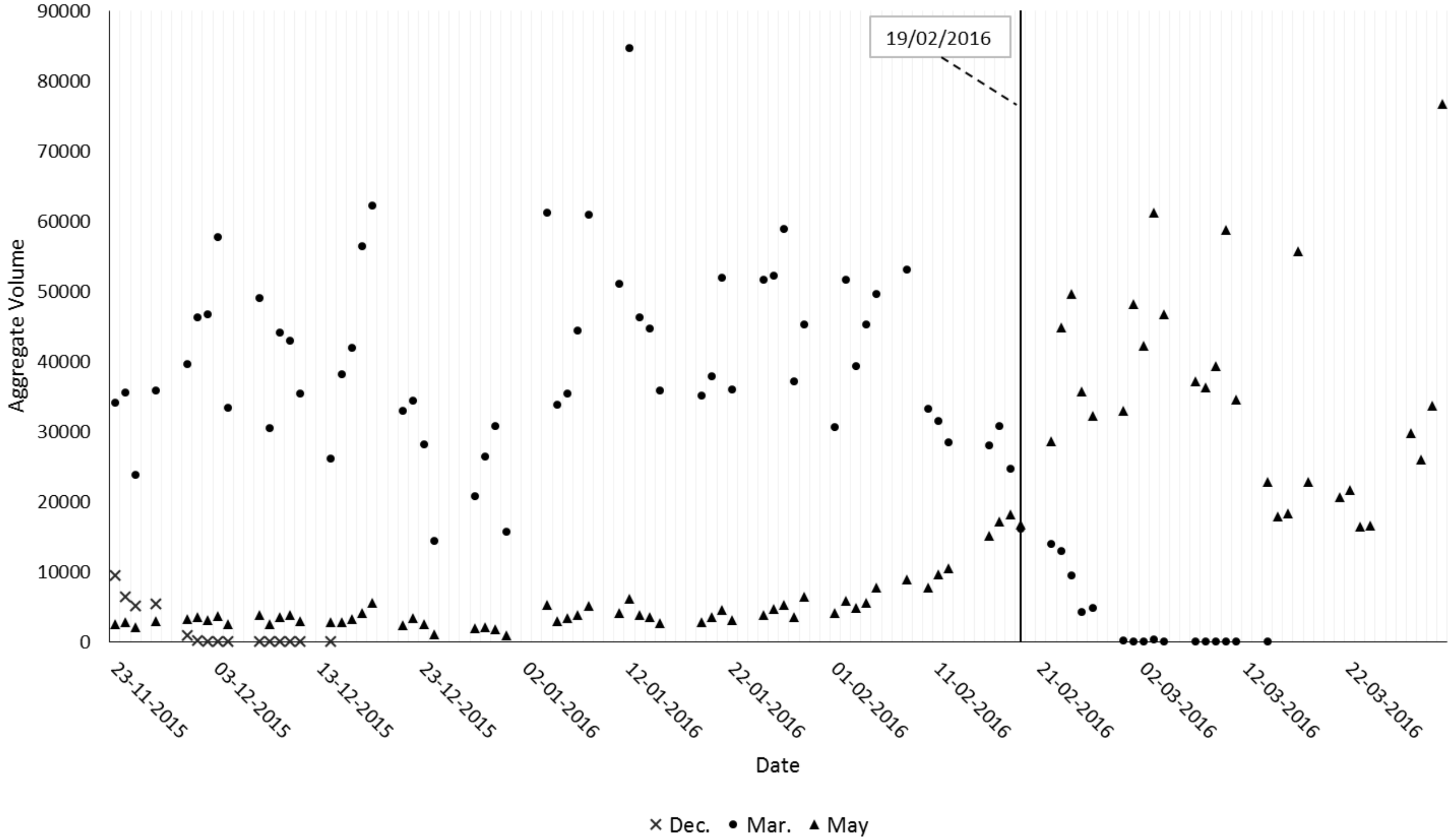

Figure 4: Roll dates for wheat

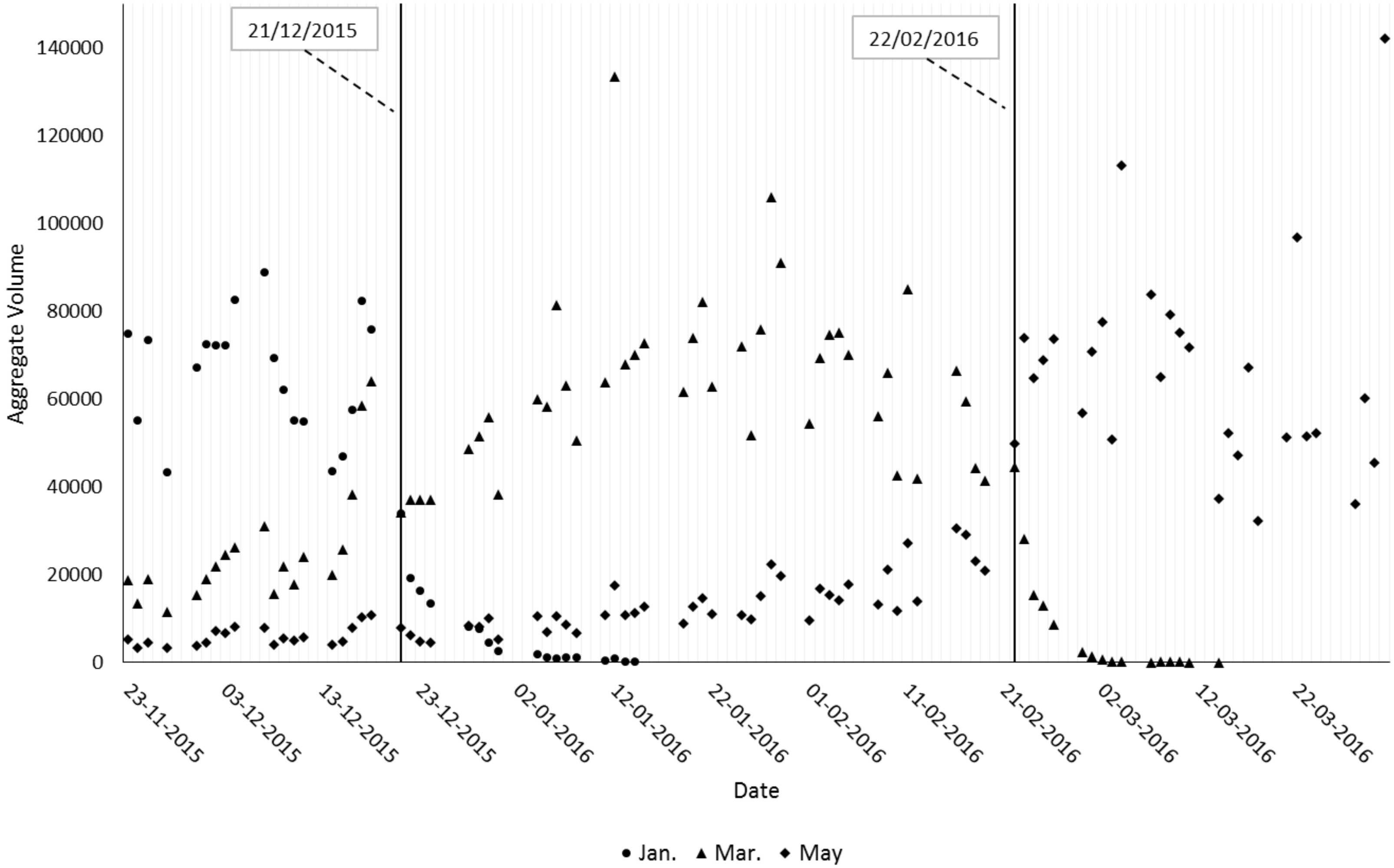

Figure 5: Roll dates for soybean

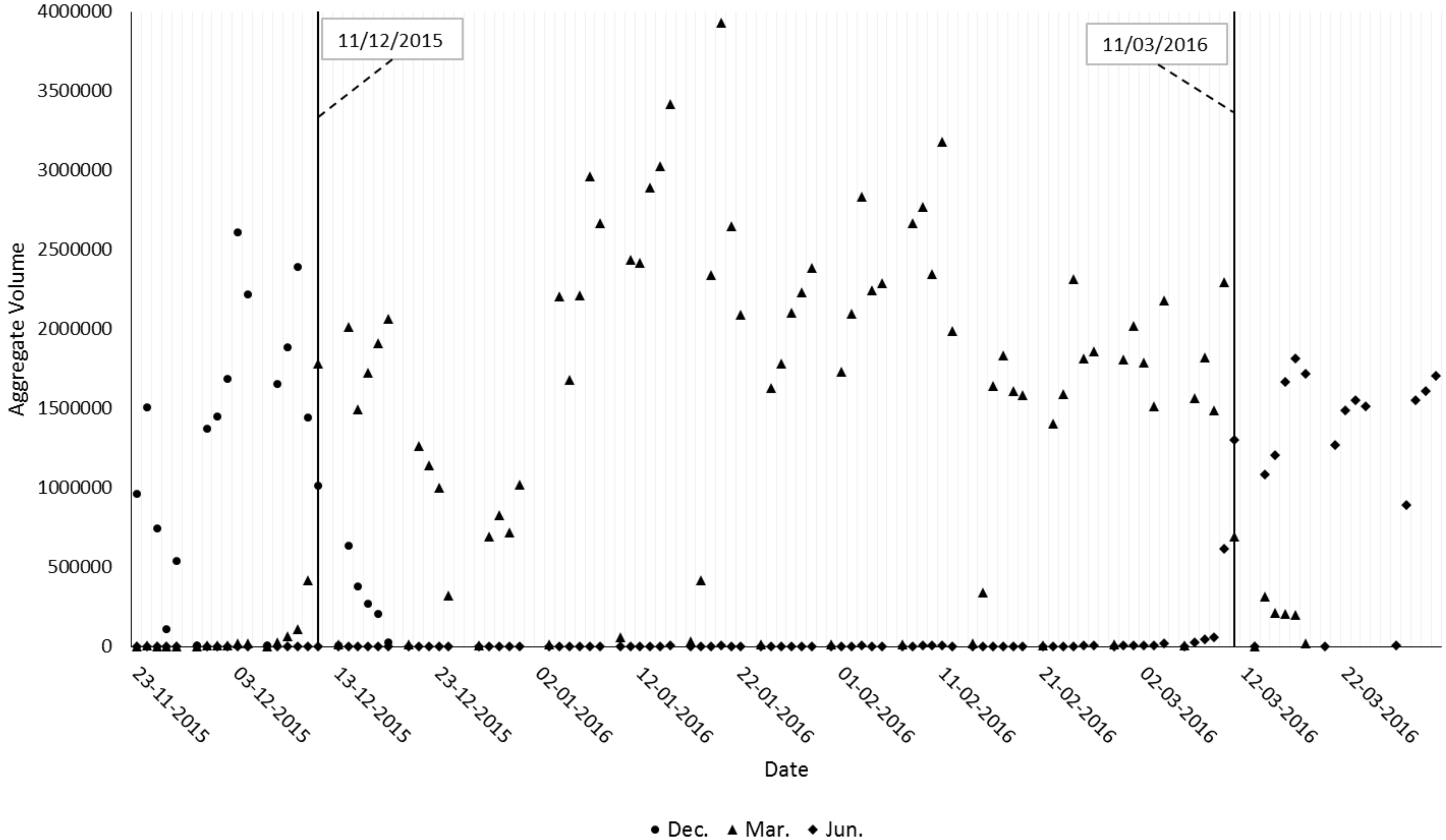

Figure 6: Roll dates for E-mini S&P 500

Table 1 summarizes the nearby contracts for our data for each commodity based on our roll date rule. Table 1 also reports the expiration date for the contracts. It can be seen in table 1 that traders of grains, on average, roll their position over three weeks before the expiration date of their contract. This roll period seems to be even longer for the livestock traders.

Reconstruction of the LOB

We use the *market depth* data files from the CME Group which provide every incremental book update required to reconstruct the LOB with Nano-second precision. Data are available to reconstruct a five-step-deep book for live cattle and lean hogs and a ten-step-deep book for corn, wheat, soybeans, and E-mini S&P 500. The data are formatted using the Financial Information eXchange (FIX/BINARY) protocol which comprises of a series of messages containing information such as bids and asks with their corresponding quantities and step in the LOB, trade prices and quantities, order sending time, and changes in the LOB such as order deletions and bids, asks and quantities updates that would define a new book. Each message is processed to reconstruct the LOB (such as in figure 7) as follows. If a message contains information on a new market order, then there is an immediate match and a trade takes place.[4] If the trade results in a partial matching of the best bid or ask, the LOB remains the same except for the change in the number of contracts at the top of the book (figure 8). On the other hand, if the trade results in a full matching of the best bid or ask, all price steps beyond the best bid or ask move one step towards the top of the book and the spread widens (figure 9). If a message contains information on a new limit order with a better price than the best bid or ask, i.e. inside the spread, the top of the book changes and the new price becomes the best bid or ask price. In this case the spread narrows and the remaining prices on the same side move one step further down along the LOB (figure 10). An order can be deleted which also updates the LOB. If it is a partial deletion, the prices in the LOB remain the same and only the corresponding quantities are altered.[5] However, if the entire quantity on a price step is cancelled or deleted, the succeeding price steps move one step upward in the LOB (figure 11).

[4] Futures trading in CME Group follows a price-time priority system, that is, orders matching the best bid or ask prices are executed first. If two orders have the same bid or ask prices, priority is given to the order that arrived first.

[5] Traders operating in CME Group have the option of submitting iceberg or hidden-size orders, which are limit orders that specify a "visible" portion of the order size. Once that quantity is filled the remaining portion of the order size is revealed. This might result in underestimation of the information contained in the LOB when the proportion of iceberg orders is high.

	Step	Price	Quantity
Ask	5	64,050	2
	4	64,000	5
	3	63,950	40
	2	63,500	1
	1	63,250	6
	Spread		
Bid	1	63,075	3
	2	62,875	27
	3	62,750	6
	4	62,700	1
	5	62,650	2

Figure 7: A five-step outright limit order book.

	Step	Price	Quantity
Ask	5	64,050	2
	4	64,000	5
	3	63,950	40
	2	63,500	1
	1	63,250	6
	Spread		
Bid	1	63,075	3
	2	62,875	27
	3	62,750	6
	4	62,700	1
	5	62,650	2

	Step	Price	Quantity
Ask	5	64,050	2
	4	64,000	5
	3	63,950	40
	2	63,500	1
	1	63,250	2
	Spread		
Bid	1	63,075	3
	2	62,875	27
	3	62,750	6
	4	62,700	1
	5	62,650	2

Figure 8: Market order arrival – buy 4 contracts at price 63,250.

	Step	Price	Quantity
Ask	5	64,050	2
	4	64,000	5
	3	63,950	40
	2	63,500	1
	1	63,250	6
		Spread	
Bid	1	63,075	3
	2	62,875	27
	3	62,750	6
	4	62,700	1
	5	62,650	2

	Step	Price	Quantity
Ask	5	64,050	2
	4	64,000	5
	3	63,950	40
	2	63,500	1
	1	63,250	6
		Spread	
Bid	1	62,875	27
	2	62,750	6
	3	62,700	1
	4	62,650	2
	5	62,600	4

Figure 9: Market order arrival – sell 3 contracts at price 63,075.

	Step	Price	Quantity
Ask	5	64,050	2
	4	64,000	5
	3	63,950	40
	2	63,500	1
	1	63,250	6
		Spread	
Bid	1	63,075	3
	2	62,875	27
	3	62,750	6
	4	62,700	1
	5	62,650	2

	Step	Price	Quantity
Ask	5	64,050	2
	4	64,000	5
	3	63,950	40
	2	63,500	1
	1	63,250	6
		Spread	
Bid	1	63,150	5
	2	63,075	3
	3	62,875	27
	4	62,750	6
	5	62,700	1

Figure 10: Limit order arrival – buy 5 contracts at price 63,150.

	Step	Price	Quantity
Ask	5	64,050	2
	4	64,000	5
	3	63,950	40
	2	63,500	1
	1	63,250	6
	Spread		
Bid	1	63,075	3
	2	62,875	27
	3	62,750	6
	4	62,700	1
	5	62,650	2

	Step	Price	Quantity
Ask	5	64,100	7
	4	64,050	2
	3	63,950	40
	2	63,500	1
	1	63,250	6
	Spread		
Bid	1	63,075	3
	2	62,875	27
	3	62,750	6
	4	62,700	1
	5	62,650	2

Figure 11: Book update message arrival – delete 5 contracts at price 64,000.

If the spread or the difference between any two steps on either buy or sell side of the book is greater than one tick (the minimum change in price allowed), traders can gain priority by submitting an order inside the spread or between two existing steps. In this case, the new price replaces the previous step and all following steps move one step down the LOB. For example, if a trader submits a buy order with a bid price higher than the third best bid price, the new bid becomes the third best bid, the previous best third bid moves to the fourth step, and, similarly, every step beyond it moves one step further from the top of the book.

The CME Group supports implied functionality which is the ability to combine spread and outright markets in one order book with the objective to increase liquidity.[6] An accurate picture of the LOB for futures contracts for a market with implied functionality at any point in time, therefore, is the one which comes from the consolidated limit order book (CLOB) that accounts for both the outright book and the implied limit order book (ILOB). The ILOB is reconstructed using data from the *market depth* files in the same way as described above for the outright book. Data are available to reconstruct a two order deep implied book for all six futures markets. The outright and the implied books are then merged into a CLOB as follows. If the price steps in the ILOB are the same as those in the LOB, the implied quantities are added to the LOB's corresponding price steps to get the CLOB. If prices are different, however, price steps coming from LOB and ILOB are compared and sorted for each bid (descending) and ask (ascending) side to form the CLOB (figure 12 and figure 13). Even though we reconstruct and employ the consolidated, CLOB, we refer to it as LOB for simplicity in what follows.

[6] An implied price is a futures order generated based on the outright market, the spread market, or other implied orders.

	Step	Price	Quantity
Ask	5	64,050	2
	4	64,000	5
	3	63,950	40
	2	63,500	1
	1	63,250	6
	Spread		
Bid	1	63,075	3
	2	62,875	27
	3	62,750	6
	4	62,700	1
	5	62,650	2

Step	Price	Quantity	
2	63,800	1	Ask
1	63,250	2	
1	63,000	2	Bid
2	62,700	4	

Figure 12: Merging LOB and ILOB.

	Step	Price	Quantity
Ask	5	64,000	5
	4	63,950	40
	3	63,800	1
	2	63,500	1
	1	63,250	8
	Spread		
Bid	1	63,075	3
	2	63,000	2
	3	62,875	27
	4	62,750	6
	5	62,700	5

Figure 13: A five-step consolidated limit order book.

Summary Statistics of the LOB

This subsection describes the main characteristics of the full LOB and its components such as transaction price, volume, all price steps and their corresponding depth at both buy and sell sides, number of book updates, and number of orders for the nearby futures contracts of, live cattle, lean hogs, corn, wheat, soybeans, and E-mini S&P 500 during Nov. 23, 2015 and Mar. 31, 2016. Table 2 reports the mean and standard deviation for all the contracts. Asks and bids and their corresponding quantities are reported for five steps for live cattle and lean hogs and ten steps for the remaining markets as disseminated by the CME Group.

The LOB updates arrive at irregular times that can be as short as a Nano second. The number of observations for each product reflects the frequency of the LOB updates. Table 2 shows that during the study time period, the LOB for E-mini S&P 500 futures contracts is updated considerably more frequently than that of agricultural commodities. Across agricultural commodities, soybeans LOB is the most dynamic book. The LOB number of updates is more or less similar for wheat and corn and for live cattle and lean hogs. Among all products, the average volume traded for corn is the highest, more than twice than that of other grains. The number of orders per trade is the highest for corn and the lowest for live cattle. On average, the BAS is about 0.046 cents (1.8 ticks) for live cattle, 0.040 cents (1.6 ticks) for lean hogs, 0.27 cent (1.1 ticks) for grains, and 25.6 cents (1 tick) for E-mini S&P 500. Along the LOB and up to the third step, corn has a considerably higher depth than the rest of the products on average, even higher than E-mini S&P 500. After the third step, E-mini S&P 500 has a higher depth than corn. Overall, the first two steps beyond the BAS seem to have a significantly higher depth than the remaining steps for agricultural commodities. Together with more or less equal price differences for bids and asks along the book, this implies that the two steps closer to the top of the book are relatively "denser" for the agricultural commodities. For E-mini S&P 500, surprisingly, further steps appear to have a slightly higher depth than the steps close to the top of the book. This means that traders of agricultural futures contracts submit more limit orders at the steps closer to the top of the book whereas traders of the E-mini S&P 500 futures contracts prefer the steps further from the top of the book. Therefore, in addition to studying differences of the LOB between agricultural commodities and E-mini S&P 500 at the aggregate level, differences of the LOB at the step level can shed light on how trading in agricultural commodity markets differs from other markets.

Price Duration

Limit and market orders that continuously update the LOB inherently arrive in an irregular timely manner. However, regularly spaced data is needed for our underlying econometric models. Previous studies suggest taking snapshots of the LOB at regular times. For example, Hasbrouck (1995) and Cao et al. (2009) both use a one-second snapshot data for thirty Dow stocks and one hundred most active Australian stocks, respectively. The time duration between snapshots is important because if it is too long, important information might be overlooked and if it is too short, we might create a data set with a lot of observations that are repeated with no new information and cause other problems such as heteroskedasticity (Engle and Russell, 1998). The literature is not clear on how to select an optimal duration for time intervals. We use the average duration of transaction price changes. Following Engle and Russell (1998), we denote every trade price change a price event and define a duration variable, d_i, given by:

$$d_i = t_i - t_{i-1} \quad [1]$$

where t_i is the time of the i^{th} transaction. We construct regularly spaced time series of the LOB for each product on the basis of how frequently their transaction price changes during the period of study. Summary statistics of the daily average durations are presented in table 3. As it can be seen, the price duration of one-second snapshots used in the finance literature (for example, Hasbrouck 1995 and Cao et al. 2009), is a good approximation for a product such as the E-mini

S&P 500 which is highly frequently traded. However, for agricultural commodities, the trading frequency and price fluctuation is considerably lower and therefore we select a longer snapshot duration to avoid a high number of repeated observations. The price duration also varies across agricultural commodities. Therefore, we choose the durations based on price events. Despite this, we repeat all our estimations and hypothesis testing for a 60-second duration for all the products to compare the outcomes. They are reported in the results section. Table 3 shows the snapshot durations for each product based on the average price durations. The table shows that the price events vary significantly across different agricultural commodities. For example, in the livestock group, live cattle average duration of 7.40 seconds is considerably lower than that of lean hogs, 11.97 seconds. This is in spite of the fact that, according to table 2, average volume for live cattle is only slightly higher than that of lean hogs and the number of contracts in all steps of the LOB on both buy and sell sides is higher for lean hogs than live cattle.[7]

Soybean futures contracts are significantly lower in volume and quantities along the LOB than those of corn according to table 2. However, average price duration is 7.60 seconds for soybeans, lower than that for corn, 8.63 seconds (table 3). Among the grains, wheat has lower volume in trade and less quantity along the LOB as well as a relatively long price duration (11.94 seconds).

Price Discovery Measures

In this section, we introduce the three measures which are used to determine the contribution of the transaction price, the spread, and the limit order book beyond the spread to price discovery of the efficient price for all six markets under study. We first present an index to capture the summary information contained in the LOB. The index is constructed following Cao et al. (2009) and it is simply a weighted price average of the bid and ask of different steps at any point in time.

Measure of the LOB Summary

An LOB consists of different price steps and the associated number of futures contracts, at any point in time. The relationship between the bid price steps and the number of contracts, related to each bid price step and aggregated across all orders, can be thought of as a market demand step function. Similarly, a market supply step function derives from the relationship between the ask price steps and the related aggregate contracts. The height of a step i in the step functions is the difference between price i and price $i-1$. For instance, the height of step 4 on the demand side is the fourth best bid less the fifth best bid. The length of a step i is the summation of the contracts across all orders for price i on each demand or supply side. The mean of the best bid and ask, denoted MID, is used to compute the first step heights for both supply and demand sides. The heights and lengths of the demand and supply step functions are, then, normalized

[7] It can be argued that an average duration based on LOB events is superior to that based on price events since the purpose of this study is to determine the informativeness of the LOB. The LOB updates, however, are greatly more frequent than the price updates and such time intervals will result in a data set with many repeated prices. Thus a duration based on LOB events must be scaled up to avoid too many repeated prices. In addition, generally when LOB updates are more frequent, so are the price updates.

using the summation of all heights and all lengths, respectively. The following weighted price reflects the price and quantity aspects of a LOB at a given point in time:

$$WP^{n_1 n_2} = \frac{\sum_{s=n_1}^{n_2}(Q_s^b P_s^b + Q_s^a P_s^a)}{\sum_{s=n_1}^{n_2}(Q_s^b + Q_s^a)},\ n_1 \leq n_2 \qquad [2]$$

where $WP^{n_1 n_2}$ is the weighted price of step n_1 to step n_2. It summarizes all the information which is contained in the LOB from step n_1 to n_2. Moreover, Q and P are quantity and price of the demand side (denoted b) or the supply side (denoted a), respectively. When $n_1 = n_2 = 1$, the weighted price becomes

$$WP^1 = \frac{Q_1^b P_1^b + Q_1^a P_1^a}{Q_1^b + Q_1^a} \qquad [3]$$

Cao et al. (2009) use MID which is the arithmetic mean of the best bid and best ask to capture the information of the spread. MID only changes when the best bid or the best ask change whereas WP^1 changes also as a result of a change in the quantities at the best bid or ask. Vo (2007) studies the quantity at the best bid and ask prices and its relationship with the BAS for Toronto Stock Exchange stocks while Frino et al. (2008) examine the relationship for three interest rates futures contracts on the Sydney Futures Exchange (SFE). The results of both studies show a negative relationship between the two variables which implies that market participants manage both price and quantity as a part of their trading strategies. Thus we use WP^1 to capture the information contained in the spread.

Table 4 and figures 14 and 19 provide the summary statistics of the shape of the average LOB for the six markets over the studied period of time. For the agricultural commodities the number of contracts resting on the second and third steps of the book, on both buy and sell sides, are considerably higher than that on the steps beyond the third step (table 4). This is comparable to the full LOB (i.e., the complete LOB before extracting the snapshots) in table 2 and suggests the possibility of a greater share of information for the first two steps of the LOB beyond the best bid and ask than the rest of the book. However, this is not the case for the E-mini S&P 500 for which the contracts appear to spread more or less equally over the second to tenth steps of the book. Moreover, in contrast to what Cao et al. (2009) observe in the Australian Stocks that the heights tend to be shorter for the steps close to the top of the book than the further away steps, the heights in our dataset are almost equal across all steps for all the products. This is also illustrated in figures 14 to 19.

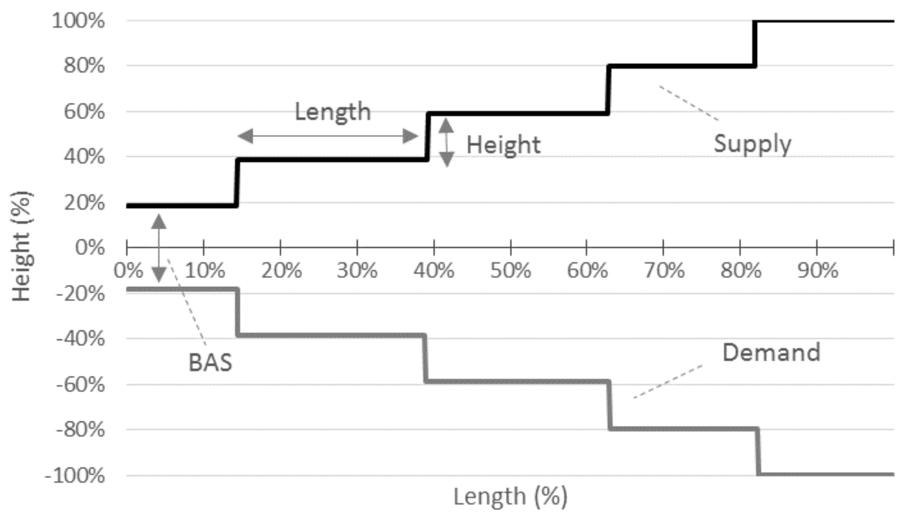

Figure 14: Live cattle average LOB.

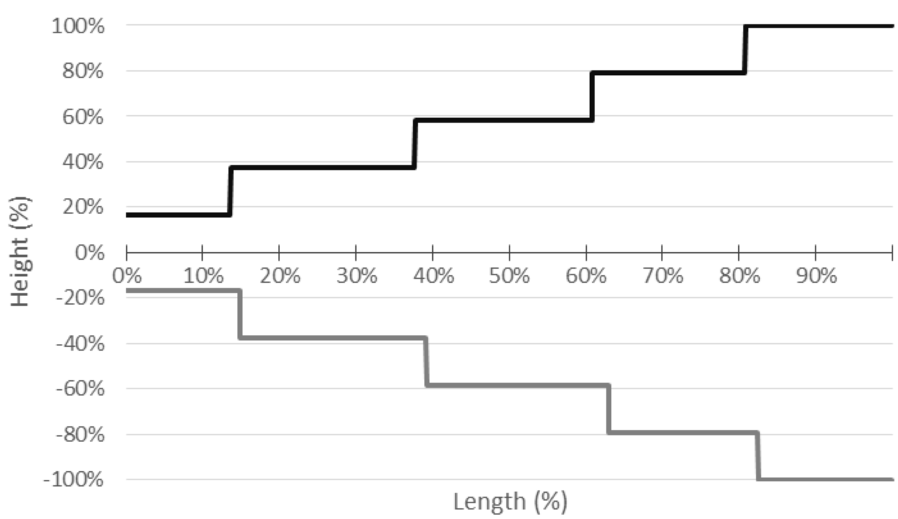

Figure 15: Lean hogs average LOB.

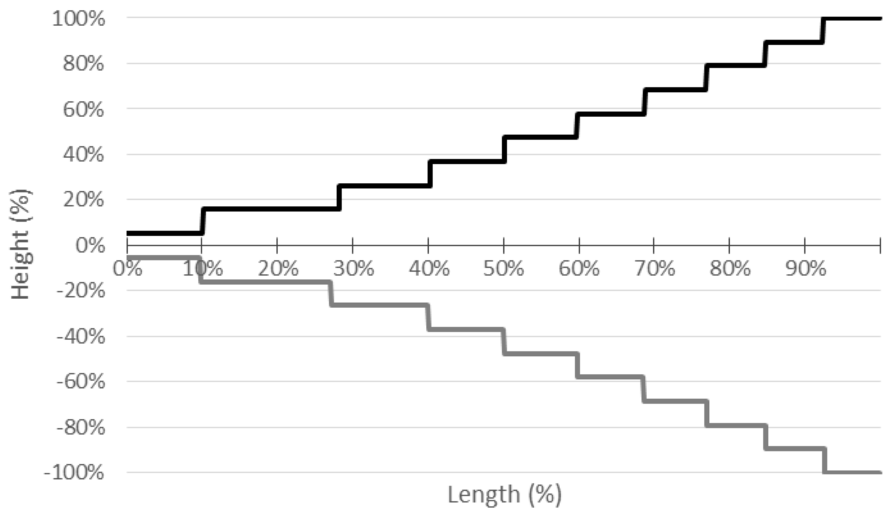

Figure 16: Corn average LOB.

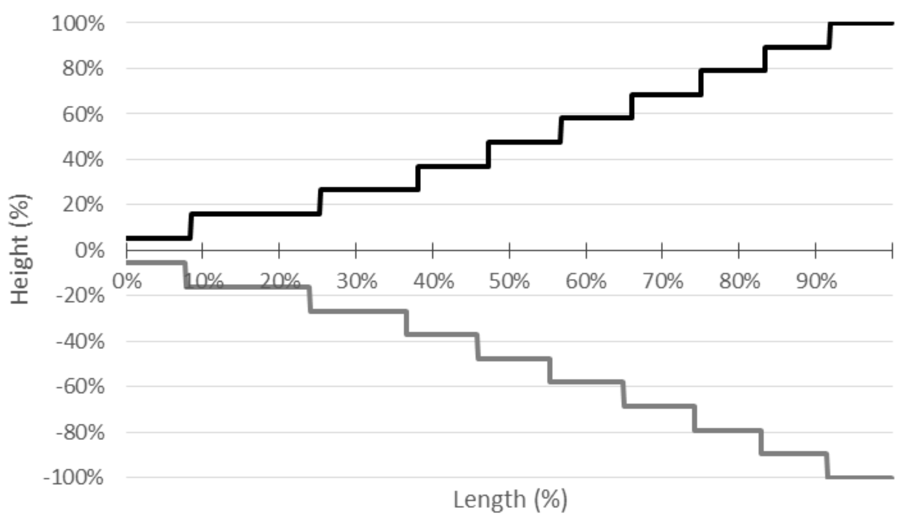

Figure 17: Wheat average LOB.

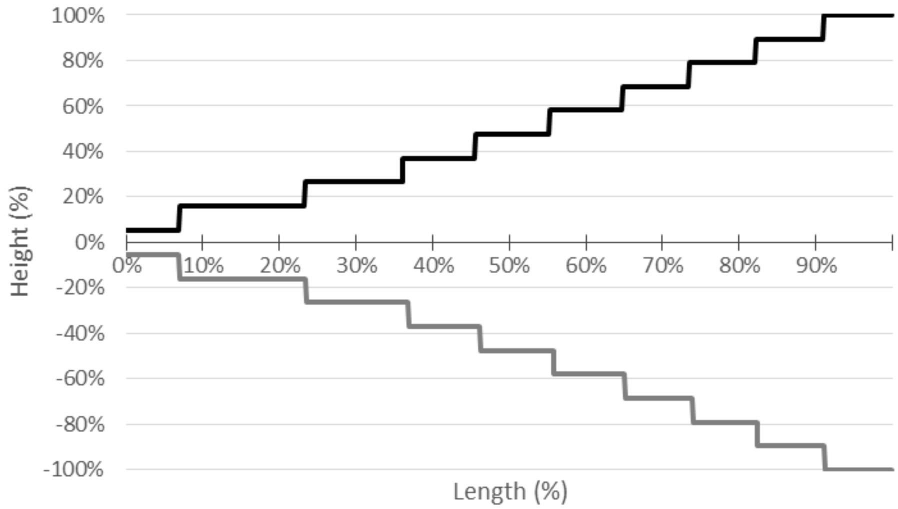

Figure 18: Soybeans average LOB.

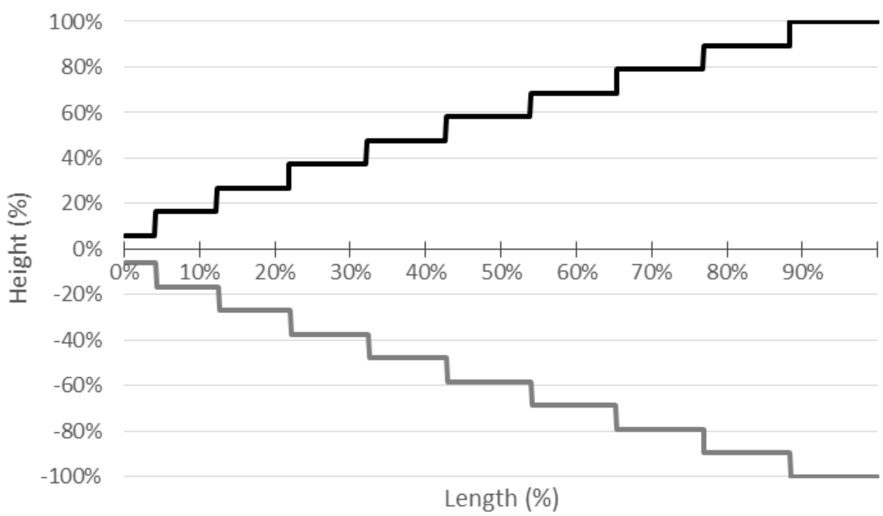

Figure 19: E-mini S&P 500 average LOB.

Error Correction Model and Measures of Information Share

Different approaches to measure the contribution of a set of price series to price discovery in a market revolve around estimating the vector error correction model (VECM). Consider K different prices of a particular commodity in the same market. A general error correction model in matrix notation for these price series can be written as:

$$\Delta \boldsymbol{Y}_t = \boldsymbol{\alpha}\boldsymbol{\beta}'\boldsymbol{Y}_{t-1} + \sum_{l=1}^{q} \boldsymbol{A}_l \Delta \boldsymbol{Y}_{t-l} + \boldsymbol{\varepsilon}_t \qquad [4]$$

where ; $\boldsymbol{Y}_t$ is the vector of price series such that $\boldsymbol{Y}_t = (y_{1t}, y_{2t}, \ldots, y_{Kt})'$, $\boldsymbol{\alpha}$ denotes the loading matrix or the matrix of coefficients which reflect how quickly price series return to their long run equilibrium, and $\boldsymbol{\varepsilon}_t$ is a white noise error term with variance-covariance matrix $\boldsymbol{\Omega}$. Prices are allowed to be functions of previous changes in their own values as well as the other price series with a matrix of coefficients of $\boldsymbol{A}$. These prices can be each non-stationary, however, they move as a group. That is, there exists a linear combination of the price series which is stationary. This means they are cointegrated and share a common stochastic trend. In equation [4], $\boldsymbol{\beta}$ represents the coefficients of this cointegration process, or the (non-unique) cointegrating matrix. Hasbrouck (1995) suggests the following form for $\boldsymbol{\beta}$:

$$\boldsymbol{\beta}'_{(K-1)\times K} = \left[\boldsymbol{\iota}_{(K-1)} \colon -\boldsymbol{I}_{(K-1)}\right] \qquad [5]$$

where $\boldsymbol{\iota}$ is a vector of 1's and $\boldsymbol{I}$ is the identity matrix.

The VECM in equation [4] can be characterized by the following vector moving average (VMA) representation (Hasbrouck, 1995):

$$\Delta \boldsymbol{Y}_t = \boldsymbol{\Psi}(L)\boldsymbol{\varepsilon}_t \qquad [6]$$

where $\boldsymbol{\Psi}(L) = \boldsymbol{\Psi_0}\boldsymbol{\varepsilon}_t + \boldsymbol{\Psi_1}\boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\Psi_2}\boldsymbol{\varepsilon}_{t-2} + \cdots$ and $\boldsymbol{\Psi_i}$ are matrices of coefficients. Equation [6] can be, alternatively, written as follows, which is known as Beveridge-Nelson decomposition (Beveridge and Nelson, 1981):

$$\boldsymbol{Y}_t = \boldsymbol{Y}_0 + \boldsymbol{\Psi}(\boldsymbol{1})\sum_{i=1}^{t} \varepsilon_i + \boldsymbol{\Psi}^*(\boldsymbol{L})\boldsymbol{\varepsilon_t} \qquad [7]$$

where $\boldsymbol{\Psi}(\boldsymbol{1})$ is the impact matrix in the lag operator, L, or the sum of the moving average coefficients. Therefore, $\boldsymbol{\Psi}(\boldsymbol{1})\varepsilon_t$ is the long run impact of an innovation, ε_t, in a price on each of the prices which is due to new information. The long run impact on all prices is the same and thus $\boldsymbol{\Psi}(\boldsymbol{1})$ has identical rows. We denote the common row in $\boldsymbol{\Psi}(\boldsymbol{1})$ by $\psi = (\psi_1, \psi_2, \ldots, \psi_K)$. The matrix $\boldsymbol{\Psi}^*(\boldsymbol{L})$**,** which is also in the lag operator, L, is the part of the price change that is resulted from transitory shocks of bid-ask bounces, inventory adjustments, or other market imperfections.

It is assumed that the price series are integrated of order one (I(1)) and that the system consists of a single common stochastic trend (Stock and Watson, 1988). That is the system has $r = K - 1$ cointegrating vectors and the impact matrix ($\boldsymbol{\Psi}(\boldsymbol{1})$) has rank one. Therefore, from the Engle-Granger representation theorem (Engle and Granger, 1987), it follows that $\boldsymbol{\beta}'\boldsymbol{\Psi}(\boldsymbol{1}) = 0$ and $\boldsymbol{\Psi}(\boldsymbol{1})\boldsymbol{\alpha} = 0$ which results in a common row in $\boldsymbol{\Psi}(\boldsymbol{1})$ or that the long-run impact of ε_t on each price is identical. Following De Jong (2002), equation [7] can be rewritten as:

$$\boldsymbol{Y}_t = \boldsymbol{Y_0} + \boldsymbol{\beta}_\perp\boldsymbol{\alpha}'_\perp\sum_{i=1}^{t} \varepsilon_i + \boldsymbol{\Psi}^*(L)\boldsymbol{\varepsilon}_t \qquad [8]$$

where $\boldsymbol{\beta}_{\perp}$ and $\boldsymbol{\alpha}_{\perp}$ are orthogonal to $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$, respectively, that is $\boldsymbol{\beta}'\boldsymbol{\beta}_{\perp} = 0$ and $\boldsymbol{\alpha}'\boldsymbol{\alpha}_{\perp} = 0$ are satisfied. Equation [8] is closely related to how Stock and Watson (1988) represent the common trend. That is, price changes have a non-stationary common factor with a permanent effect (f_t) and a stationary transient component (G_t) given by:

$$\boldsymbol{Y}_t = \boldsymbol{f}_t + \boldsymbol{G}_t \qquad [9]$$

The common trend representation in Stock and Watson (1988) represented in equation [9] and the Beveridge-Nelson decomposition of equation [7] are the basis for the information share measures which follow.

1. Granger and Gonzalo Permanent-Transitory Effect (PT)

Gonzalo and Granger (1995) suggest that each of the prices in a system potentially contributes to the common trend or the efficient price. Therefore, the common factor is defined as a combination of prices given by:

$$f_t = \boldsymbol{\Gamma}\boldsymbol{Y}_t \qquad [10]$$

where $\boldsymbol{\Gamma}$ is a $1 \times K$ vector of coefficients for the common factor with elements $(\gamma_1, \gamma_2, \dots, \gamma_K)$. Under this specification of the common trend, the error correction term is not allowed to Granger cause the common factor in the long run. They show that $\boldsymbol{\Gamma}$ is orthogonal to the vector of the error correction coefficients $\boldsymbol{\alpha}$ and the common trend representation, therefore, can be written as:

$$f_t = \boldsymbol{\alpha}'_{\perp}\boldsymbol{Y}_t \qquad [11]$$

Finally, the PT measure of the contribution of the j^{th} price to the efficient price is related to γ_j in $\boldsymbol{\Gamma}$ or $\alpha_{\perp j}$ in $\boldsymbol{\alpha}_{\perp}$. That is, the PT information share only depends on γ or $\alpha_{\perp}$. Harris et al. (2002) normalize the vector coefficients of the common trend such that the sum of the price information shares equals one. Based on Harris et al. (2002), PT can be computed using:

$$PT_j = \frac{\alpha_{\perp j}}{\sum_{j=1}^{K} \alpha_{\perp j}} \qquad [12]$$

2. Hasbrouck Information Share (IS)

Hasbrouck (1995) also uses the common factor representation by Stock and Watson (1988) to develop an information share in order to measure the contribution of different market prices to the efficient price discovery. The difference between IS and the previous approach is that in IS the variance of the common factor is decomposed and each price contributes to the efficient price based on how its variance contributes to the variance of the efficient price. The variance of the common factor innovations is given by

$$var(\psi\varepsilon_t) = \psi\boldsymbol{\Omega}\psi' \qquad [13]$$

where ψ is a common row vector in the $\boldsymbol{\Psi}(\mathbf{1})$ matrix in [7]. We compute the parameters in $\boldsymbol{\Psi}(\mathbf{1})$ directly using Johansen's factorization and the estimation coefficients from the VECM in equation [4] by:

$$\boldsymbol{\Psi}(\mathbf{1}) = \boldsymbol{\beta}_{\perp}\boldsymbol{\Pi}\boldsymbol{\alpha}'_{\perp} \qquad [14]$$

where

$$\boldsymbol{\Pi} = \left(\boldsymbol{\alpha}'_{\perp}(\boldsymbol{I} - \sum_{l=1}^{q}\boldsymbol{A}_l)\boldsymbol{\beta}_{\perp}\right)^{-1} \qquad [15]$$

and $\boldsymbol{I_K}$ is the identity matrix.

The IS of the j^{th} price, then, can be calculated by

$$IS_j = \frac{\left(\psi_j\sigma_j\right)^2}{\psi\boldsymbol{\Omega}\psi'} \qquad [16]$$

where σ_j is the j^{th} price's standard deviation in the variance-covariance matrix $\boldsymbol{\Omega}$.

Hasbrouck (1995) suggests that if the variance-covariance matrix of residuals ($\boldsymbol{\Omega}$) in the VECM representation (equation [4]) is not diagonal, that is the price innovations are significantly correlated across the price series, IS and PT can result in misleading information shares. Hasbrouck (1995) uses the Cholesky factorization of the residuals covariance matrix to eliminate the contemporaneous correlation. Based on the Cholesky factorization,

$$\boldsymbol{\Omega} = \boldsymbol{MM}' \qquad [17]$$

where $\boldsymbol{M}$ is a lower triangle matrix. The IS measure can then be written as,

$$IS_j = \frac{[\psi\boldsymbol{M}]_j{}^2}{\psi\boldsymbol{\Omega}\psi'} \qquad [18]$$

Even though the IS calculated using equation [18] solves the correlation problem, it creates another issue that is the measure being sensitive to the ordering of prices in the system. This occurs because when correlation exists, that is the nondiagonal elements of M are nonzero, the IS measure imposes more weights on the prices that appear earlier in the system. To overcome this problem, Hasbrouck (1995) proposes calculation of upper and lower bounds for each price by placing them first and last in the system. In the multivariate cases, all permutations of the variables must be computed to find the upper and lower bounds (Hasbrouck, 2002). [8]

Baillie et al. (2002), De Jong (2002), and Yan and Zivot (2010) show that the PT measure can be computed by:

$$PT_j = \frac{\psi_j}{\sum_{j=1}^{k} \psi_j} \qquad [19]$$

Therefore, we use the coefficients of the long run impact matrix, $\boldsymbol{\Psi}(\mathbf{1})$, computed using equation [14], to measure IS and PT in the equations [18] and [19], respectively.

3. Modified Information Share (MIS) and other Information Share Metrics

Several studies have been carried out to address the drawbacks of PT and IS metrics or to extend them for more general settings. Among them Yan and Zivot (2010) show that the aforementioned measures can be misleading if different price series have different levels of noise. Different levels of noise in price series of a common asset can arise if, for example, the minimum tick size differs in two different markets trading the asset, different inventory management in the markets, bid–ask bounce, or other microstructure frictions and market imperfections (Yan and Zivot, 2010; Putnins, 2013). Yan and Zivot (2010) argue that only the IS measure can provide information on the relative informativeness of individual price series, however, the IS measure for a series may be higher due to more information contained in that series relative to other series or it can be higher if the other series are more noisy even though the former is not necessarily containing more information. Moreover, PT which can be computed using elements of the error correction coefficient vector α in equation [12] (Baillie et al., 2002), measures the way in which prices adjust to lagged differences in their transitory components. In the case of two price series (Yan and Zivot, 2010; Putnins, 2013), PT of price series 1 reflects how sensitive price series 2 is, relative to price series 1, to lagged transitory shocks and vice versa. Yan and Zivot (2010) show that PT in fact reflects the dynamic responses of price series to the transitory shocks and not the permanent shocks. However, IS reflects how price series respond to both permanent and transitory shocks. Yan and Zivot (2010) propose combining IS and PT to specifically measure impounding of new information to account for, on the one hand, the information content of each series, and on the other hand, control for differences in the noise

---

[8] Baillie et al. (2002) show how PT and IS can result in similar information shares in the absence of contemporaneous correlation in the residuals and under the assumption that there exists only one single common factor in the system. They show that both IS and PT measures are essentially derived from $\boldsymbol{\alpha}_{\perp}$. This is true because under a single common factor, the coefficients of $\boldsymbol{\Psi}(\mathbf{1})$ are different from those of $\boldsymbol{\alpha}_{\perp}$ only by a scalar which cancels out when calculating the information share. They measure IS using the following formulas for when a price is placed first (upper bound) and last (lower bound) in the system, respectively by $IS_{j1} = \frac{\left[\sum_{j=1}^{K} \gamma_j m_{j1}\right]^2}{\boldsymbol{\Gamma\Omega\Gamma'}}$ and $IS_{jK} = \frac{(\gamma_K m_{KK})^2}{\boldsymbol{\Gamma\Omega\Gamma'}}$ where m_{ij} is the element in the $\boldsymbol{M}$ matrix which is on the i^{th} row and j^{th} column and γ_j are directly estimated from the VECM model in equation [4] without deriving the VMA representation. Note that in their IS measure, the common row elements of $\boldsymbol{\Gamma}$ i.e. $(\gamma_1, \gamma_2, \dots, \gamma_K)$ are used instead of those of $\boldsymbol{\Psi}(\mathbf{1})$.

levels. That is, when there are only two price series in the system, combining the two measures can rule out the transitory shock effect on the price series to capture the long run effect of a permanent shock only. Putnins (2013) tests the measure further and derive the informational leadership (IL) measure. This approach, however, is more appealing in bivariate systems and how successful IL measure is in ruling out the effect of the transitory shocks when there are more than two price series is unknown (Putnins, 2013).

The modified information share (MIS) was developed by Lien and Shrestha (2009) to address the problems with the previous approaches, i.e. a potential nondiagonal variance covariance matrix of residuals in PT and the order sensitivity in IS. In MIS, on the one hand, a type of factorization is used to eliminate potential contemporaneous correlations. On the other hand, instead of the variance-covariance matrix Ω which is sensitive to the ordering of price series, MIS uses the innovation correlation matrix which is invariant to different orders of price series in the system. The factorization procedure is as follows. Consider the innovation correlation matrix denoted by $\boldsymbol{\Phi}$ of the VEC model in [4]. Consider also a diagonal matrix $\boldsymbol{\Lambda}$ where its elements are the eigenvalues of the correlation matrix ($\boldsymbol{\Phi}$). And the corresponding eigenvectors are given by the columns of matrix $\boldsymbol{G}$. Finally, suppose matrix $\mathbf{V}$ is diagonal and its elements are the standard deviation of the innovations or the square roots of diagonal elements of $\boldsymbol{\Omega}$. Lien and Shrestha (2009) show that the innovations can be transferred to the following:

$$\varepsilon_t = \boldsymbol{F} z_t \tag{20}$$

where z_t is the transferred innovation with zero mean, $E[z_t] = 0$, and identity covariance matrix, $E[z_t z_t{}'] = \boldsymbol{I}$. Moreover,

$$\boldsymbol{F} = \left[\boldsymbol{G}\boldsymbol{\Lambda}^{-1/2}\boldsymbol{G}'\mathbf{V}^{-1}\right]^{-1} \tag{21}$$

and $\boldsymbol{\Omega} = \boldsymbol{F}\boldsymbol{F}'$. This factorization results in the MIS measure given by

$$MIS_j = \frac{[\psi \boldsymbol{F}]_j{}^2}{[\psi \boldsymbol{F}]\boldsymbol{\Omega}[\psi \boldsymbol{F}]'} \tag{22}$$

MIS has the desirable feature that it calculates an information share which is order invariant and it accounts for contemporaneous correlations in innovations. This is particularly important when correlation exists in the residuals. When correlation is nearly nonexistent, MIS approaches IS and PT, and when innovations are completely correlated MIS approaches $\frac{1}{K}$ or an equal share for all the K price series. Lien and Shrestha (2014) extended their MIS measure by placing the diagonal coefficients of $\boldsymbol{\Gamma}$ in the cointegration matrix, $\boldsymbol{\beta}$ (equation [4]), instead of 1's and using the cointegration matrix in the form of $\boldsymbol{\beta}'_{(K-1)\times K} = \left[\boldsymbol{\iota}_{(K-1)}\colon -\boldsymbol{\Gamma}_{(K-1)}\right]$ instead of $\boldsymbol{\beta}'_{(K-1)\times K} = \left[\boldsymbol{\iota}_{(K-1)}\colon -\boldsymbol{I}_{(K-1)}\right]$, before the factorization of innovations variance-covariance matrix using the matrix of correlations. They suggest that the generalization is suitable for the cases where the price discovery contribution of different but related financial securities are analyzed such as price discovery in markets for different security issued by the same firm and call the new measure Generalized Information Share (GIS).

Another measure of price discovery (let us denote it by TLS) was developed by Grammig and Peter (2013) to address the IS problem of non-uniqueness, particularly for longer sampling intervals. They assume a multivariate mixture distribution to develop a measure which they call "tail-dependent information shares". Like MIS, TLS follows from Hasbrouck's (1995) contribution of a price series variance to the variance of the efficient price as the measure of the series information share by means of reduced VECM long run impact coefficients. However, the variance decomposition under TLS is performed using a VECM which is extended by the mixture parameters and estimated by a two-step process. This, unlike IS, results in an order neutral measure and is claimed to be superior to IS and PT when correlations of price innovations in the tails differ from those in the center of the distributions.

Lien and Wang (2016) conduct a research using simulated data to compare the IS upper and lower bound midpoint with the two unique, more recent, information shares of MIS and TLS. They find that TLS performs poorly for the simulated data even when the underlying assumptions of the approach are met. Moreover, their results show that MIS at most marginally improves the information share computed by the IS midpoint. They, therefore, support the use of the IS midpoint as a method of computing the information shares of different price series.

We use the two common information share metrics, i.e. Hasbrouck (1995) IS and Gonzalo and Granger (1995) PT, as well as Lien and Shrestha (2009) MIS metric to determine the contribution of the BAS and the LOB beyond the BAS to price discovery of the six markets under study. An extended survey on the developments in information share research and the studies using different approaches in various settings can be found in Narayan and Smyth (2015).

There is very little work done on the agricultural commodities futures market and little information available regarding the contribution of each step of a limit order book to price discovery for these commodities. We compute the three introduced measures of information share to have a better understanding of the role of a limit order book in price discovery in the agricultural commodities futures market. To make our results comparable to other markets, we also calculate the measures for the popular E-mini S&P 500. We also break down the steps beyond BAS into two price series to assess whether the information contained in the LOB is uniformly distributed or the steps closer to the BAS contain different amount of information relative to the steps farther away. The analysis aims to improve our understanding of the price discovery process in futures markets, and in particular in agricultural commodity markets.

Results

We estimated two models one with three variables (where $\boldsymbol{Y}_t = (Price \;\; WP^1 \;\; WP^{2-5})'$ for live cattle and lean hogs and $\boldsymbol{Y}_t = (Price \;\; WP^1 \;\; WP^{2-10})'$ for the other markets) and one with four variables (where $\boldsymbol{Y}_t = (Price \;\; WP^1 \;\; WP^{2-3} \;\; WP^{4-5})'$ for live cattle and lean hogs, $\boldsymbol{Y}_t = (Price \;\; WP^1 \;\; WP^{2-3} \;\; WP^{4-10})'$ for corn, wheat, and soybeans, and $\boldsymbol{Y}_t = (Price \;\; WP^1 \;\; WP^{2-5} \;\; WP^{6-10})'$ for the E-mini S&P 500). The presence of a unit root at levels and first differences for each price series is tested using the Augmented Dickey–Fuller unit root test and all price series are found to be I(1) for most days. The unit root test is performed using different number of lags up to eighty with no significant change in the results. The VEC model in equation [4] is estimated for each day and the average information share computed using equations [18], [19], and [22], over all days is reported for each product. The cointegration rank is tested using three statistics namely Johansen's trace statistic at the 99% confidence level, minimizing the Schwarz Bayesian information criterion (BIC), and minimizing the Hannan and Quinn information criterion (HQIC) (Gonzalo and Pitarakis, 1998; Aznar and Salvador, 2002). The rank is found to be 2 for the three-price series model and 3 for the four-price series model for almost all days. Therefore, for most days, the assumption that the system has $r = K - 1$ cointegrating vectors and the impact matrix, $\mathbf{\Psi(1)}$, has rank one is satisfied. Moreover, all models are estimated using eighty lags for each price series.[9] Lagrange Multiplier (LM) test and autocorrelation functions were used to test for autocorrelation in the residuals and showed no significant autocorrelation in the residuals.

Tables 5 to 10 report means of each of the price discovery measures (i.e. IS, PT, and MIS) for all futures contracts for the three-variable model. The information shares computed for the days for which the model underlying assumptions are not fully met, I(1) price series or a single cointegration vector present in the model, are in line with the information shares for the rest of the days and thus are included in the calculation of the average information share over all days. Tables 5-10 show the information share measures for the model with three price series i.e. Price, WP^{1}, and $WP^{2\text{-}5}$ (for meats) and $WP^{2\text{-}10}$ (for grains and the E-mini S&P 500). The three measures expectedly lead to different information shares, however, with an overall consistency. Daily information share measures show that for days when one measure of information share for a price series is high (low), the other two information shares also result in higher (lower) shares. For the majority of the weeks, the information share for the LOB beyond the BAS follows a V-shape pattern for E-mini S&P 500. This means that a week starts with a high information share for the LOB beyond the BAS. The information share declines for the following weekdays and reaches a minimum in the middle of the week, mainly on Wednesdays. The information share of the LOB beyond the BAS then increases reaching previous high levels at the end of trading weeks. For the agricultural commodities this pattern is less apparent meaning that the information contained in the steps beyond the BAS can be higher or lower for any weekday during a particular week.

Based on the aggregate results in tables 5 to 10, the three measures indicate a substantial contribution of the steps beyond the BAS to price discovery for all futures contracts higher than that of Cao et al. (2009) study for the Australian stocks (22%). Considering the midpoint IS measure (IS_M), the contribution of the steps beyond the BAS for grains to price discovery is about 30% (29.49% for corn, 30.16% for soybeans, and 31.01% for wheat). IS_M suggests a slightly lower share for the steps beyond the BAS for lean hogs (26.69%), E-mini S&P 500 (26.92%), and live cattle (27.29%). The share of the steps beyond the BAS are considerably higher according to PT. The information share contained in the steps beyond the BAS based on PT, are 32.76% and 35.21% for lean hogs and live cattle, respectively. PT indicates a high

[9] The number of lags is usually determined by minimizing a criterion such as AIC or BIC. These tests were performed using a maximum number of 80 lags. Using 80 lags results in minimum AIC and BIC and the ACF of the residuals show no significant autocorrelation. Thus, 80 lags were added to the model. Also 80 lags cover over ten minutes of time for the agricultural commodities datasets used here given that our time intervals are greater than 7 seconds for all agricultural commodities. Moreover, Lehecka et al. (2014) show that announcement effects of major USDA reports on corn futures prices and volume disappear in ten minutes, using intraday Chicago Board of Trade data for July 2009 to May 2012.

information share for the steps beyond the BAS for wheat (42.26%) and soybeans (43.72%). The information share for the steps further from the top of the book is even higher for corn (52.90%) and E-mini S&P 500 (59.44%). MIS results are similar to PT information share for the steps beyond the BAS for lean hogs, live cattle, wheat, and soybeans. MIS of the steps beyond the BAS for live cattle is 34.92% and is 33.25% for lean hogs. For grains, MIS indicates a 40.10% share for the steps beyond the BAS for wheat, 40.66% for soybeans, and 42.49% for corn. For E-mini S&P 500, the MIS for the steps beyond the BAS is 45.19%.

A similar share for the steps beyond the BAS for live cattle and lean hogs is found despite the fact that during the period of study, the average price event duration is significantly lower for live cattle relative to lean hogs (7 compared to 12 seconds in table 3). This implies that even though the live cattle futures market is considerably more active, that is with more price changes than lean hogs, traders in both markets utilize the steps of the LOB beyond the BAS in their trading strategies to a similar extent. Moreover, according to table 3, lean hogs and wheat have approximately the same duration (12 seconds) but the share for the steps beyond the BAS is higher for wheat than lean hogs. In addition, live cattle duration is lower than grains while the information share for the steps beyond the BAS for grains is higher than that of live cattle. A comparison among the agricultural commodities, shows that live cattle and soybeans are the commodities with the highest number of price events whereas lean hogs and wheat have the highest price durations. Moreover, the informativeness of the LOB beyond the BAS is similar for lean hogs and live cattle and is similar for soybeans and wheat and corn. This suggests a higher level of noise in the live cattle futures market relative to the rest of the agricultural commodities.

The futures markets analyzed appear to be more homogenous with regards to the information contained in the BAS. With the exception of PT for E-mini S&P 500, the BAS contributes to the price discovery about 30%. The BAS contains more information in price discovery for the meats group compared to the other products, according to the measures. Price contains a moderate contribution to price discovery for the grains and E-mini S&P 500 but higher information for live cattle and lean hogs (tables 5 - 10).

Tables 5 to 10 also show that the higher-lower bound for IS is considerably wide for all products. This emphasizes the importance of a measure for price discovery which is order insensitive. As a robustness check, we calculated the information share measures for all the markets using 60-second interval data. The results are reported in the appendix, table A1. The information share measures obtained from the 60-second interval data is particularly different for E-mini S&P 500 relative to our main results. Using 60-second interval data for E-mini S&P 500 increases the contribution of the BAS and decreases the information share measures for the steps beyond the BAS. This can be explained by the large number of book updates that are missed which come mainly from the steps beyond the BAS and may carry information. For this reason, the difference between the measures derived from the main data and those derived from the data with 60-second intervals for lean hogs and wheat is small. And the measures become increasingly different for the products for which the LOB is more frequently updated such as corn.

We also calculated the information share measure using MID instead of WP^1 to compare the results with those of Cao et al. (2009). The results are reported in the appendix, table A2. The results of the IS midpoint metric for E-mini S&P 500 when MID is used instead of WP^1 are similar to Cao et al. (2009) results for IS midpoint for the one hundred active Australian stocks. They find the information share of the LOB for nine steps beyond the BAS to be 22.47% on average using the midpoint Hasbrouck (1995) information share. Moreover, they find that Price has a share of 23.15% and MID's share is 54.50%. Our normalized IS midpoint results for the E-mini S&P 500 are 24.27% for Price, 50.17% for MID, and 25.56% for $WP^{2\text{-}10}$ (table A2).[10] Comparing our results of the model with WP^1 in tables 5 - 10 and that of the model with MID in table A2, the information share of MID appears to be considerably higher than that of WP^1 according to all three measures for all the products studied. This may be due to the lower level of noise in MID than WP^1 and not necessarily a reflection of a higher level of informativeness (Yan and Zivot, 2010; Putnins, 2013). MID is only sensitive to changes in the best bid and best ask whereas WP^1 fluctuates with a change in the corresponding quantities of best bid and ask as well. This also implies that a considerable part of the BAS updates that comes from the change in quantities at the best bid and ask is due to trades that carry noise and are not necessarily due to an inflow of information in the market. Moreover, impatient and liquidity traders submit their orders primarily at the best prevailing bid and ask (Pascual and Veredas, 2009).

By visualizing the shape of the average LOB for the six products in table 4 and figures 14 - 19, we observe a special shape for agricultural commodities' LOB. The number of contracts on the second and third steps seem to be considerably higher than steps further from the BAS. To study this further, we break down the steps of the LOB beyond the BAS into two variables, one with steps closer to the top of the book and one farther, to examine whether the information of the steps beyond the BAS is uniformly distributed. In order words, we examine if informed traders use steps which are closer to the top of the book rather than the steps further from the BAS to exploit their private information. For E-mini S&P 500, the contracts spread more or less equally across all steps. Therefore, in order to assess the information contained in the LOB in more detail, we construct two variables for the agricultural commodities (one for the second and third steps and one for the remaining steps) and two variables for the E-mini S&P 500 (one for the second to fifth steps and one for the remaining steps). The aggregate information share measures are reported in table 11. Similar to the three-variable model, the information share measures produce different shares but are consistent across all days. The information share for the variables that are relatively lower (higher) based on one measure is lower (higher) for to the other two measures as well. An important pattern again is evident for the information share of the steps beyond the BAS for E-mini S&P 500. The V-shape pattern of the whole LOB beyond the BAS that was observed in the three-variable model is due to the steps closer to the BAS. And the information share of the steps closer to the BAS and the steps further away along the book have opposite patterns. In other words, when a trading week starts, the steps closer to the BAS contain higher information which decreases on the following days and increases by the end of the week before the weekend. However, the steps farther start the week with lower information, increase in information over the next weekdays and end the week with low information once again. The higher information contained in the steps closer to the top of the book offset the opposite behaviour of the steps farther and overall gives the LOB information share a V-shape pattern for each week. For agricultural commodities, the pattern of the information share appears to be more random.

[10] The IS midpoint measure for different price series does not necessarily sum to a hundred percent. Thus, the IS midpoint for the price series is normalized to add up to a hundred percent.

Table 11 shows that the information share measures indicate a higher share for the steps closer to the top of the book for most of the products. From the two meats, live cattle steps closer to the BAS have relatively more information share based on the three metrics results. For lean hogs, IS_M, PT, and MIS of WP^{2-3} are, respectively, 18.80, 26.05, and 26.48 percent, comparable to 17.06, 17.06, and 18.77 percent of WP^{4-5}. The difference is more pronounced for live cattle. IS, PT, and MIS of WP^{2-3} for live cattle are, respectively, 19.83, 28.79, and 28.26 percent, as opposed to 16.04, 17.14, and 18.64 percent of WP^{4-5}. This means that in live cattle futures markets, despite the high number of price events and the possibility of high noise, the noise is not necessarily concentrated in the steps close to the top of the book. For grains, the measures generate mixed results but it seems that the two steps beyond the BAS are as informative as the seven steps further away. For corn, all measures suggest a slightly higher share for WP^{2-3} than WP^{4-10} whereas for soybeans the opposite is true. For E-mini S&P 500, WP^{2-5} has a significantly higher information share than WP^{6-10}, according to all three measures.

Cao et al. (2009) also examined the distribution of information across the LOB steps beyond the BAS using two variables for the 100 most active Australian stocks. They divided their LOB steps beyond the BAS into two variables one from step 2 to 4 and the other from step 5 to 10. Despite considerable higher depth in the steps closer to the BAS, they find that the contribution of the steps further from the BAS to price discovery is higher than the steps closer to the BAS. They link this phenomenon to the existence of two types of orders in the steps at or near the top of the book: orders that are used in searching for hidden orders or orders that are "faked" and canceled in a few seconds and are used to "spoof" the opposite side of the market in order to improve the price levels in the trader's favour. This "fishing" and "spoofing" may be less present in the steps further from the top of the book and thus those steps may be less noisy and more informative. We find that in the futures markets, this is not the case and the steps closer to the top of the book contain more information than the steps further from the BAS. This can be an indication that faked and hidden orders are used to a lesser degree in the futures markets. Across the agricultural commodities, soybeans might be the market with relatively higher noise present near the top of the book as opposed to corn and meats.

Conclusions

Price discovery is a fundamental function of futures markets and is defined as the incorporation of information to prices through the actions of traders. Recent finance literature has found evidence that, as a part of their trading strategies, informed traders may submit limit orders instead of market orders for a variety of reasons such as minimizing their market impact or to avoid disclosing their private information to other traders. If so, the BAS and the steps of LOB beyond BAS contain valuable information and contribute to price discovery of the underlying asset. This is the first attempt to examine the informativeness of the LOB beyond BAS for agricultural commodities.

We reconstruct the LOB for five major agricultural commodities namely live cattle, lean hogs, corn, wheat, and soybean as well as a very active futures contract namely E-mini S&P 500 using all book updates over the period of Nov. 23, 2015 to Mar. 31, 2016. We find that there is large number of contracts existing on the bids and asks beyond the best bid and ask for all the products. For agricultural commodities, most contracts exist at steps two and three compared to the BAS and the steps farther from the top of the book. For E-mini S&P 500, we find that the contracts are uniformly distributed along the LOB steps.

Three information share metrics are applied to assess whether informed traders use the steps of the LOB beyond the BAS in futures markets and especially in agricultural commodity markets in their trading strategies. Our results show a substantial contribution of the steps of the LOB beyond the BAS to price discovery in futures markets, higher than what the literature finds for stock markets (i.e. Cao et al., 2009). Across the products studied, IS indicates a 27% share for the contribution of the steps beyond the BAS to price discovery for meats and E-mini S&P 500 and about 29%-31% for the grains. PT information share suggests that the contribution of the LOB steps beyond the BAS is in the 30%-40% range for meats, the 40%-50% range for soybeans and wheat, and the 50%-60% range for corn and E-mini S&P 500. MIS information share for the LOB steps beyond the BAS is about 33%-35% for the meats, 40%-42% for the grains, and 45% for E-mini S&P 500. Moreover, the information contained in the BAS and Price is higher for the meats than the rest of the products studied.

Along the LOB steps, our results show that the steps closer to the top of the book have more information relative to the steps farther from the best bid and ask. This contrasts the finding by Cao et al. (2009) and suggests that faked and spoofing trades are less present in the futures markets relative to the stock markets. Considering the daily estimates of the information share metrics, our results show that the steps beyond the BAS but closer to the top of the book contain more information during the early and late weekdays and less information in the middle of the week, mainly Wednesdays creating a V-shaped pattern for the LOB information share for E-mini S&P 500. On the other hand, the steps beyond the BAS but farther from the top of the book have the opposite behaviour and start the week with less information, the information increases and reaches its peak in the middle of the week, and finishes the week with lower information. This means that informed traders in the E-mini S&P 500 futures markets use the steps closer to the top of the book more on the early and late days of the week and the steps farther in the middle of the week. The behaviour of the LOB steps for agricultural commodities appears to be less systematic.

Table 1: Nearby contracts for all the futures contracts

	Contracts – Expiration	Days before rolling
	February – 29 Feb. 2016	23 Nov. 2015 – 18 Jan. 2016
Live Cattle	April – 30 Apr. 2016	19 Jan. 2016 – 16 Mar. 2016
	June – 30 Jun. 2016	17 Mar. 2016 – 31 Mar. 2016
	February – 12 Feb. 2016	23 Nov. 2015 – 20 Jan. 2015
Lean Hogs	April – 14 Apr. 2016	21 Jan. 2016 – 22 Mar. 2016
	June – 14 Jun. 2016	23 Mar. 2016 – 31 Mar. 2016
Corn	March – 14 Mar. 2016	23 Nov. 2015 – 21 Feb. 2016
	May – 13 May 2016	22 Feb. 2016 – 31 Mar. 2016
	January – 14 Jan. 2016	23 Nov. 2015 – 20 Dec. 2015
Soybeans	March – 14 Mar. 2016	21 Dec. 2015 – 21 Feb. 2016
	May – 13 May 2016	22 Feb. 2016 – 31 Mar. 2016
Wheat	March – 14 Mar. 2016	23 Nov. 2015 – 18 Feb. 2016
	May – 13 May 2016	19 Feb. 2016 – 31 Mar. 2016
E-min S&P 500	December – 18 Dec. 2015	23 Nov. 2015 – 10 Dec. 2015
	March – 18 Mar. 2016	11 Dec. 2015 – 10 Mar. 2016
	June – 17 Jun. 2016	11 Mar. 2016 – 31 Mar. 2016

Table 2: Mean and Standard Deviation of the Full LOB

	Lean Hogs	Live Cattle	Corn	Wheat	Soybeans	E-mini S&P 500
Observations[(1)]	6,176,794	6,961,908	25,305,597	21,285,007	42,625,431	386,421,232
Price[(2)]	66835.90 (6666.22)	132277.90 (4498.3)	366.09 (7.23)	471.54 (11.44)	880.52 (14.79)	195664.50 (7608.52)
Volume (contract)[(3)]	(2.06) (2.82)	2.19 (2.78)	12.07 (32.25)	4.72 (9.19)	5.32 (10.32)	10.13 (25.99)
Number of Orders[(4)]	2.08 (1.06)	2.03 (0.98)	5.78 (8.01)	3.69 (3.77)	4.06 (4.31)	4.86 (7.64)
Bid 1	66816.08 (6667.63)	132255.10 (4499.03)	365.97 (7.23)	471.41 (11.44)	880.39 (14.79)	195651.70 (7608.56)
Ask 1	66856.39 (6664.52)	132300.90 (4497.35)	366.23 (7.23)	471.68 (11.44)	880.66 (14.79)	195677.30 (7608.50)
Bid 2	66790.74 (6667.82)	132229.50 (4499.21)	365.72 (7.23)	471.16 (11.44)	880.14 (14.79)	195626.70 (7608.56)
Ask 2	66881.76 (6664.30)	132326.60 (4497.20)	366.48 (7.23)	471.93 (11.44)	880.91 (14.79)	195702.30 (7608.50)
Bid 3	66765.45 (6667.99)	132203.80 (4499.40)	365.47 (7.23)	470.91 (11.44)	879.89 (14.79)	195601.70 (7608.56)
Ask 3	66907.11 (6664.14)	132352.20 (4497.12)	366.73 (7.23)	472.18 (11.44)	881.16 (14.79)	195727.30 (7608.50)
Bid 4	66740.16 (6668.19)	132178.20 (4499.57)	365.22 (7.23)	470.66 (11.44)	879.64 (14.79)	195576.70 (7608.56)
Ask 4	66932.39 (6664.02)	132377.70 (4497.04)	366.98 (7.23)	472.43 (11.44)	881.41 (14.79)	195752.30 (7608.50)
Bid 5	66714.76 (6668.41)	132152.40 (4499.83)	364.97 (7.23)	470.41 (11.44)	879.39 (14.79)	195551.70 (7608.56)
Ask 5	66957.75 (6663.83)	132403.30 (4496.99)	367.23 (7.23)	472.68 (11.44)	881.66 (14.79)	195777.30 (7608.50)
Bid 6			364.72 (7.23)	470.16 (11.44)	879.14 (14.79)	195526.70 (7608.56)
Ask 6			367.48 (7.23)	472.93 (11.44)	881.91 (14.79)	195802.30 (7608.50)

Table 2 continued

	Live Cattle	Lean Hogs	Corn	Wheat	Soybeans	E-mini S&P 500
Bid 7			364.47	469.91	878.89	195501.70
			(7.23)	(11.44)	(14.79)	(7608.56)
Ask 7			367.73	473.18	882.16	195827.30
			(7.23)	(11.44)	(14.79)	(7608.50)
Bid 8			364.22	469.66	878.64	195476.70
			(7.23)	(11.44)	(14.79)	(7608.56)
Ask 8			367.98	473.43	882.41	195852.30
			(7.23)	(11.44)	(14.79)	(7608.50)
Bid 9			363.97	469.41	878.39	195451.70
			(7.23)	(11.44)	(14.79)	(7608.56)
Ask 9			368.23	473.68	882.66	195877.30
			(7.23)	(11.44)	(14.79)	(7608.50)
Bid 10			363.72	469.16	878.14	195426.70
			(7.23)	(11.44)	(14.79)	(7608.56)
Ask 10			368.48	473.93	882.91	195902.30
			(7.23)	(11.44)	(14.79)	(7608.50)
Quant.1 buy	5.91	6.33	377.69	45.26	62.45	160.35
(contract)	(7.80)	(8.67)	(431.07)	(47.32)	(66.55)	(139.51)
Quant.1 sell	5.56	6.48	375.27	43.92	64.00	159.31
(contract)	(6.99)	(10.66)	(414.80)	(47.31)	(68.54)	(140.88)
Quant.2 buy	10.44	11.12	716.52	98.47	147.00	314.07
(contract)	(10.97)	(10.95)	(467.46)	(61.80)	(90.48)	(163.47)
Quant.2 sell	10.16	12.40	682.75	95.02	150.04	313.77
(contract)	(10.19)	(15.27)	(408.51)	(61.09)	(101.81)	(169.32)
Quant.3 buy	10.47	11.09	538.92	79.01	119.31	365.46
(contract)	(11.06)	(10.95)	(379.75)	(64.52)	(86.04)	(178.83)
Quant.4 buy	8.95	9.48	419.09	62.15	88.90	388.33
(contract)	(11.19)	(11.60)	(313.07)	(64.76)	(71.83)	(187.62)
Quant.4 sell	8.28	10.86	387.64	59.10	93.46	389.95
(contract)	(10.71)	(16.77)	(279.37)	(62.31)	(95.33)	(193.41)

Table 2 continued

	Live Cattle	Lean Hogs	Corn	Wheat	Soybeans	E-mini S&P 500
Quant.5 buy (contract)	8.46 (12.33)	8.72 (12.06)	401.01 (306.12)	63.44 (73.51)	90.66 (79.50)	402.50 (191.85)
Quant.5 sell (contract)	7.98 (12.19)	10.90 (19.80)	371.72 (285.90)	59.79 (62.09)	98.29 (109.71)	404.79 (199.81)
Quant.6 buy (contract)			369.64 (284.56)	63.41 (74.08)	88.44 (82.84)	416.85 (194.19)
Quant.6 sell (contract)			350.73 (301.49)	59.39 (63.00)	95.88 (111.69)	419.54 (202.98)
Quant.7 buy (contract)			343.95 (276.23)	61.49 (72.41)	83.53 (83.83)	423.42 (194.46)
Quant.7 sell (contract)			321.98 (305.87)	57.89 (65.23)	89.62 (109.35)	426.40 (202.03)
Quant.8 buy (contract)			323.65 (261.35)	57.76 (70.05)	80.67 (82.89)	427.13 (193.24)
Quant.8 sell (contract)			310.74 (297.76)	54.98 (65.41)	86.58 (102.85)	430.61 (200.79)
Quant.9 buy (contract)			312.04 (260.15)	55.74 (70.88)	82.27 (89.39)	424.55 (189.48)
Quant.9 sell (contract)			304.45 (308.39)	54.56 (68.26)	89.06 (107.74)	429.22 (202.42)
Quant.10 buy (contract)			302.25 (249.49)	56.05 (74.46)	83.82 (92.80)	426.65 (187.98)
Quant.10 sell (contract)			298.57 (303.45)	54.53 (69.90)	91.66 (115.24)	433.34 (207.77)

Notes: Numbers between parentheses are standard deviations

Ask (Bid) x is the mean ask (bid) price in the x^{th} step of the LOB across all reconstructed LOBs in the sample period. Quant. x is the mean number of contracts in the x^{th} step of the LOB.

(1) Total number of LOB updates, or the total number of books for the nearby contracts.

(2) Average transaction price for the nearby contracts. The price unit is 0.001 cent per pound for livestock, cent per bushel for grains, and cent per contract for E-mini S&P 500. Bid and Ask units are the same as that for the transaction price.

(3) Average volume per transaction for the nearby contracts.

(4) Average number of orders per transaction for the nearby contracts.

Table 3: Price Event Durations

	Average price event duration (seconds)	Standard Deviation	Minimum price event duration (seconds)	Maximum price event duration (seconds)
Live Cattle	7.40	2.43	2.08	14.60
Lean Hogs	11.97	3.72	6.68	26.51
Corn	8.63	2.51	2.48	16.51
Wheat	11.94	4.07	3.87	25.70
Soybeans	7.60	2.12	1.84	14.70
E-mini S&P	1.12	0.65	0.35	4.12

Note: Price event is defined as a transaction price change, and price event duration is the time between two price changes. The average price event durations for each product is the mean of each day price durations averaged across all days. Minimum and maximum price event are the minimum and maximum average durations for all days in the sample period.

Table 4: Summary statistics of the shape of the average sample LOB, Nov. 23, 2015 – Mar. 31, 2016

	Length (%)		Height (%)		Length (number of contracts)		Height (cents)	
Steps	Buy	Sell	Buy	Sell	Buy	Sell	Buy	Sell
Live Cattle								
1	14.39	14.25	18.22	18.28	6.20	5.74	23.51	23.51
2	24.52	24.98	20.38	20.44	10.95	10.45	27.52	26.25
3	23.96	23.60	20.40	20.42	10.80	10.08	25.81	25.53
4	19.43	19.12	20.45	20.41	8.95	8.28	26.25	25.42
5	17.70	18.06	20.55	20.46	8.52	8.13	26.83	25.52
Number of observations: 176,142; mean price: 132498; mean volume: 1.74; mean number of orders: 1.95								
Lean Hogs								
1	14.87	13.57	16.78	16.79	6.44	6.29	21.07	21.07
2	24.37	24.11	20.75	20.73	10.78	11.70	25.86	25.26
3	23.77	23.11	20.74	20.88	10.73	11.55	25.20	26.30
4	19.46	20.02	20.88	20.78	9.02	10.39	26.27	25.38
5	17.52	19.19	20.84	20.81	8.34	10.73	25.50	25.42
Number of observations: 116,082; mean price: 66282.95; mean volume: 1.62; mean number of orders: 1.99								
Corn								
1	9.38	10.12	5.32	5.30	407.20	401.17	0.13	0.13
2	17.72	18.09	10.54	10.52	760.14	712.05	0.25	0.25
3	12.88	12.09	10.54	10.52	557.94	477.20	0.25	0.25
4	10.04	9.85	10.51	10.50	443.41	401.16	0.25	0.25
5	9.75	9.57	10.51	10.59	429.93	392.65	0.25	0.25
6	8.84	8.96	10.52	10.52	390.11	372.34	0.25	0.25
7	8.41	8.20	10.51	10.52	366.82	334.70	0.25	0.25
8	7.78	7.84	10.51	10.52	335.03	320.76	0.25	0.25
9	7.78	7.71	10.51	10.50	328.54	319.11	0.25	0.25
10	7.41	7.56	10.51	10.50	315.21	312.85	0.25	0.25
Number of observations: 175,463; mean price: 365.90; mean volume: 4.93; mean number of orders: 3.98								

Table 4 continued

	Length (%)		Height (%)		Length (number of contracts)		Height (cents)	
Steps	Buy	Sell	Buy	Sell	Buy	Sell	Buy	Sell
				Wheat				
1	7.79	8.36	5.57	5.54	47.94	47.22	0.13	0.13
2	16.25	16.90	10.53	10.50	103.35	97.98	0.25	0.25
3	12.54	12.81	10.49	10.53	81.31	77.12	0.25	0.25
4	9.33	9.26	10.49	10.50	64.54	59.07	0.25	0.25
5	9.42	9.34	10.51	10.48	68.04	61.48	0.25	0.25
6	9.56	9.34	10.48	10.53	68.79	61.83	0.25	0.25
7	9.36	9.01	10.48	10.50	66.76	59.76	0.25	0.25
8	8.66	8.36	10.48	10.46	61.68	56.32	0.25	0.25
9	8.55	8.37	10.49	10.49	60.67	56.76	0.25	0.25
10	8.55	8.25	10.49	10.46	60.84	56.10	0.25	0.25
Number of observations: 126,789; Price mean: 471.69; mean volume: 2.92; mean number of orders: 3.03								
				Soybeans				
1	6.90	6.92	5.48	5.47	63.73	64.19	0.13	0.13
2	16.58	16.34	10.50	10.59	153.24	154.03	0.25	0.26
3	13.32	12.88	10.52	10.50	123.06	121.91	0.25	0.25
4	9.41	9.36	10.50	10.51	91.42	94.19	0.25	0.25
5	9.62	9.75	10.50	10.52	95.19	101.99	0.25	0.25
6	9.26	9.42	10.52	10.51	93.57	98.70	0.25	0.25
7	8.79	8.77	10.50	10.48	89.37	91.13	0.25	0.25
8	8.55	8.61	10.50	10.48	85.59	88.46	0.25	0.25
9	8.67	8.91	10.48	10.48	88.09	92.24	0.25	0.25
10	8.90	9.03	10.50	10.48	90.04	93.51	0.25	0.25
Number of observations: 201,377; mean price: 879.08; mean volume: 3.40; mean number of orders: 3.35								

Table 4 continued

	Length (%)		Height (%)		Length (number of contracts)		Height (cents)	
Steps	Buy	Sell	Buy	Sell	Buy	Sell	Buy	Sell
E-mini S&P 500								
1	4.27	4.08	5.71	5.72	158.94	157.85	14.05	14.05
2	8.31	8.18	10.68	10.65	311.12	311.61	25.91	25.66
3	9.53	9.67	10.47	10.44	356.51	359.31	25.11	25.00
4	10.39	10.14	10.44	10.47	378.33	382.38	25.01	25.11
5	10.46	10.58	10.53	10.44	392.24	397.86	25.28	25.00
6	11.15	11.22	10.43	10.44	407.05	414.61	25.00	25.00
7	11.26	11.51	10.43	10.44	412.87	420.85	25.00	25.00
8	11.59	11.40	10.43	10.44	417.89	425.06	25.00	25.00
9	11.44	11.53	10.43	10.44	413.79	422.93	25.00	25.00
10	11.61	11.68	10.43	10.52	414.48	426.24	25.00	25.27

Number of observations: 2,263,729; mean price: 197,931.20; mean volume: 4.42; mean number of orders: 3.03

Table 5: Summary Statistics of the Average Information Share Measures; Live Cattle (%)

		Price	WP^1	$WP^{2\text{-}5}$
	Mean	77.67	63.18	55.14
	Median	76.18	60.48	51.11
IS_H	Sd	10.99	17.27	18.19
	Min	57.70	29.79	29.62
	Max	98.05	96.52	95.33
	Mean	11.69	8.66	5.36
	Median	7.59	5.23	2.84
IS_L	Sd	11.53	8.46	5.95
	Min	0.11	0.00	0.00
	Max	43.46	31.56	24.32
IS_M	Mean(N)	**40.31**	**32.40**	**27.29**
	Mean	**26.83**	**37.96**	**35.21**
	Median	20.69	31.36	32.61
PT	Sd	20.27	22.07	23.51
	Min	2.13	0.49	0.14
	Max	81.41	86.97	82.28
	Mean	**28.54**	**36.53**	**34.92**
	Median	23.73	32.31	34.16
MIS	Sd	13.34	12.36	13.25
	Min	12.32	10.49	13.23
	Max	71.30	63.00	62.72

Note: The information share metrics in the table are estimated for each day and averaged over all days in the sample period. IS_M is calculated by averaging the IS_H and IS_L for each day and then averaging over all days. IS_M for the variables does not necessarily add up to 100 percent. Thus, we normalized IS_M for the three variables to add up to 100 percent, and its mean is referred to as Mean(N), to be comparable to the other measures.

Table 6: Summary Statistics of the Average Information Share Measures; Lean Hogs (%)

		P	WP^1	WP^{2-5}
	Mean	79.59	58.55	52.79
	Median	79.46	54.70	49.63
IS_H	Sd	11.37	18.24	16.15
	Min	44.45	30.37	30.93
	Max	100.00	99.28	95.74
	Mean	13.80	8.08	5.47
	Median	11.22	3.92	3.11
IS_L	Sd	11.07	8.88	5.89
	Min	0.04	0.00	0.00
	Max	44.26	42.07	21.11
IS_M	Mean(N)	**42.78**	**30.53**	**26.69**
	Mean	**32.20**	**35.04**	**32.76**
	Median	27.34	31.69	29.19
PT	Sd	21.06	23.23	20.81
	Min	1.25	0.18	0.66
	Max	98.73	91.30	86.54
	Mean	**31.49**	**35.26**	**33.25**
	Median	28.30	33.35	31.40
MIS	Sd	13.58	13.55	12.70
	Min	10.12	11.51	11.97
	Max	75.28	71.31	62.61

Table 7: Summary Statistics of the Average Information Share Measures; Corn (%)

		P	WP^1	WP^{2-10}
	Mean	53.84	62.87	50.38
	Median	49.01	64.76	46.14
IS_H	Sd	20.75	17.87	23.38
	Min	17.30	18.22	8.66
	Max	96.87	96.98	95.83
	Mean	16.87	16.09	12.21
	Median	11.24	13.21	9.68
IS_L	Sd	17.12	15.11	10.02
	Min	0.00	0.11	0.00
	Max	75.03	60.13	40.57
IS_M	Mean(N)	**33.31**	**37.20**	**29.49**
	Mean	**20.87**	**26.23**	**52.90**
	Median	18.98	24.50	55.40
PT	Sd	15.80	17.95	21.99
	Min	0.08	1.98	1.33
	Max	74.78	84.72	91.13
	Mean	**24.91**	**32.60**	**42.49**
	Median	20.86	31.90	39.95
MIS	Sd	18.17	14.41	16.11
	Min	2.26	7.70	9.61
	Max	82.69	71.35	76.48

Table 8: Summary Statistics of the Average Information Share Measures; Wheat (%)

		P	WP^{1}	WP^{2-10}
	Mean	68.69	59.75	58.41
	Median	69.67	58.60	57.70
IS_H	Sd	15.00	14.97	16.36
	Min	33.82	30.75	26.41
	Max	99.45	93.82	94.32
	Mean	12.75	8.98	9.08
	Median	10.18	7.36	7.54
IS_L	Sd	11.09	8.11	7.42
	Min	0.00	0.00	0.00
	Max	52.38	38.65	29.30
IS_M	Mean(N)	**37.42**	**31.58**	**31.01**
	Mean	**28.52**	**29.22**	**42.26**
	Median	26.06	27.29	42.98
PT	Sd	18.19	16.61	20.14
	Min	0.58	0.12	0.07
	Max	89.51	81.54	81.11
	Mean	**27.49**	**32.42**	**40.10**
	Median	25.61	32.46	40.31
MIS	Sd	14.13	10.78	13.18
	Min	7.44	11.27	11.56
	Max	74.15	63.73	66.17

Table 9: Summary Statistics of the Average Information Share Measures; Soybeans (%)

		P	WP^1	WP^{2-10}
	Mean	66.92	64.50	57.60
	Median	66.28	63.62	56.93
IS_H	Sd	13.97	13.08	17.85
	Min	31.18	34.06	26.18
	Max	99.60	92.38	99.39
	Mean	11.44	10.22	8.49
	Median	8.01	8.63	6.94
IS_L	Sd	10.27	8.25	7.84
	Min	0.05	0.03	0.02
	Max	48.08	32.80	34.75
IS_M	Mean(N)	**35.75**	**34.09**	**30.16**
	Mean	**26.12**	**30.16**	**43.72**
	Median	21.73	30.71	42.93
PT	Sd	16.96	16.48	22.09
	Min	1.71	1.22	4.44
	Max	83.32	68.22	95.08
	Mean	**25.48**	**33.86**	**40.66**
	Median	22.42	33.51	40.69
MIS	Sd	13.06	10.38	14.01
	Min	4.60	13.45	16.39
	Max	65.60	57.46	72.39

Table 10: Summary Statistics of the Average Information Share Measures; E-mini S&P 500 (%)

		P	WP^1	WP^{2-10}
	Mean	77.30	74.06	58.00
	Median	81.75	74.10	59.17
IS_H	Sd	14.23	12.69	9.95
	Min	42.29	49.77	26.52
	Max	96.01	95.93	79.63
	Mean	10.16	6.26	3.81
	Median	8.92	3.35	3.87
IS_L	Sd	7.68	7.03	1.56
	Min	0.01	0.00	0.02
	Max	31.79	30.28	8.21
IS_M	Mean(N)	**38.09**	**34.98**	**26.92**
	Mean	**23.56**	**17.00**	**59.44**
	Median	23.30	15.39	63.46
PT	Sd	12.37	10.76	14.04
	Min	0.71	0.49	4.17
	Max	51.11	57.22	81.11
	Mean	**22.59**	**32.22**	**45.19**
	Median	22.22	31.73	47.44
MIS	Sd	10.06	5.82	6.43
	Min	4.98	21.98	25.99
	Max	45.01	45.56	53.32

Table 11: The Information Share Measures for Four-Variable Model (%)

		IS_H	IS_L	IS_M(N)	PT	MIS
Live Cattle	Price	79.65	8.66	37.53	23.92	24.20
	WP^1	57.12	5.45	26.60	30.15	28.90
	WP^{2-3}	42.91	3.75	19.83	28.79	28.26
	WP^{4-5}	34.45	3.30	16.04	17.14	18.64
Lean Hogs	Price	80.80	9.83	39.10	28.43	26.36
	WP^1	53.01	5.05	25.04	28.46	28.40
	WP^{2-3}	40.30	3.29	18.80	26.05	26.48
	WP^{4-5}	35.70	3.83	17.06	17.06	18.77
Corn	Price	51.37	11.01	28.08	15.69	18.17
	WP^1	61.68	9.95	32.25	19.06	25.09
	WP^{2-3}	36.88	8.00	20.21	34.34	29.61
	WP^{4-10}	35.59	7.65	19.47	30.91	27.13
Wheat	Price	70.01	8.52	34.12	22.80	21.52
	WP^1	59.22	6.62	28.61	24.43	26.56
	WP^{2-3}	33.06	3.40	15.84	27.16	25.77
	WP^{4-10}	43.84	5.46	21.42	25.38	26.14
Soybeans	Price	65.81	6.44	31.21	18.42	18.23
	WP^1	63.68	7.32	30.67	25.04	27.28
	WP^{2-3}	33.31	4.00	16.12	28.20	26.28
	WP^{4-10}	44.94	6.02	22.01	28.34	28.21
E-mini S&P	Price	73.23	4.46	31.61	14.48	13.51
	WP^1	82.00	8.19	36.69	21.05	28.57
	WP^{2-5}	44.26	2.04	18.84	45.85	34.46
	WP^{6-10}	30.72	0.90	12.86	18.62	23.46

Note: The measures in the table were calculated for each day and averaged over all days. IS_M is calculated by averaging the IS_H and IS_L for each day and then averaging over all days. IS_M for the variables does not necessarily add up to 100 percent. Thus, we normalized IS_M for the three variables, denoted by IS_M(N), to add to 100 percent to be comparable to the other measures.

Appendix 1 – Information Share Measures for 60-second snapshot and MID cases.

Table A1: The Information Share Measures (%) – 60 second snapshots

		IS_L	IS_H	IS_M	PT	MIS
	Price	16.60	95.65	56.12	32.02	32.77
Live Cattle	WP^1	1.40	52.68	26.93	33.73	33.46
	$WP^{2\text{-}5}$	1.29	53.09	27.19	34.25	33.77
	Price	16.28	93.29	54.78	32.43	32.52
Lean Hogs	WP^1	2.12	50.44	26.16	33.60	33.75
	$WP^{2\text{-}5}$	2.33	53.61	27.97	33.97	33.72
	Price	15.08	79.95	47.52	26.69	29.78
Corn	WP^1	6.52	59.74	33.02	30.63	32.82
	$WP^{2\text{-}10}$	6.35	51.62	28.98	42.68	37.41
	Price	12.19	87.01	49.60	27.10	29.52
Wheat	WP^1	2.81	49.69	26.20	29.24	31.46
	$WP^{2\text{-}10}$	5.50	62.78	34.14	43.66	39.02
	Price	12.35	89.69	51.02	25.49	29.39
Soybeans	WP^1	2.97	53.28	28.17	32.57	33.43
	$WP^{2\text{-}10}$	3.42	58.91	31.16	41.94	37.18
	Price	16.44	97.67	57.06	23.07	31.12
E-mini S&P	WP^1	0.99	57.38	29.06	40.06	34.61
	$WP^{2\text{-}10}$	0.29	48.42	24.36	36.87	34.27

Table A2: The Information Share Measures for the model with MID (%)

		IS_L	IS_H	IS_M	PT	MIS
	Price	6.84	75.19	41.02	17.24	23.59
Live Cattle	MID	9.18	75.38	42.28	51.40	41.43
	$WP^{2\text{-}5}$	3.18	61.98	32.58	31.36	34.98
	Price	9.88	78.97	44.42	23.76	27.36
Lean Hogs	MID	6.90	65.94	36.42	44.55	38.55
	$WP^{2\text{-}5}$	4.19	61.40	32.80	31.69	34.27
	Price	8.73	43.50	26.11	12.26	15.95
Corn	MID	19.49	77.16	48.32	39.71	40.81
	$WP^{2\text{-}10}$	8.44	58.48	33.46	48.02	43.24
	Price	8.22	64.64	36.43	20.30	22.35
Wheat	MID	9.95	70.25	40.10	41.45	38.38
	$WP^{2\text{-}10}$	6.72	63.75	35.23	38.25	39.27
	Price	5.10	59.67	32.39	14.87	17.78
Soybeans	MID	12.74	77.45	45.10	45.63	41.50
	$WP^{2\text{-}10}$	6.05	65.74	35.90	39.50	40.72
	Price	2.86	52.30	27.58	11.34	14.42
E-mini S&P	MID	21.67	92.34	57.01	61.96	45.87
	$WP^{2\text{-}10}$	0.61	57.46	29.04	26.70	39.34

References

Aidov A., 2013. “Three Essays on Market Depth in Futures Markets.” PhD dissertation, Florida International University. FIU Electronic Theses and Dissertations. Paper 974. Available via http://digitalcommons.fiu.edu/etd/974 (last accessed on 24/01/2016).

Aidov, A., and R.T. Daigler. 2015. “Depth Characteristics for the Electronic Futures Limit Order Book.” *Journal of Futures Markets* 35(6):542–560.

Anand, A., and A. Subrahmanyam. 2008. “Information and the Intermediary: Are Market Intermediaries Informed Traders in Electronic Markets?” *The Journal of Financial and Quantitative Analysis* 43(1):1–28.

Anand, A., V.A. Gatchev, L. Madureira, C.A. Pirinsky, and S. Underwood. 2001. “Geographic Proximity and Price Discovery: Evidence from NASDAQ.” *Journal of Financial Markets* 14(2):193–226.

Aznar, A., and M. Salvador. 2002. “Selecting the Rank of the Cointegration Space and the Form of the Intercept Using an Information Criterion.” *Econometric Theory* 18:926–947.

Bailliea, R.T, G.G. Bootha, Y. Tseb, and T. Zabotina. 2002. “Price Discovery and Common Factor Models.” *Journal of Financial Markets* 5(3):309–321.

Baruch, S. 2005. “Who Benefits from an Open Limit-Order Book?” *Journal of Business* 78:1267–1306.

Bauwens, L., and P. Giot. 2001. “Econometric Modelling of Stock Market Intraday Activity.” The Netherlands: Kluwer Academic Publishers.

Beveridge, S., C.R. and Nelson. 1981. “A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the ‘Business Cycle’.” *Journal of Monetary Economics* 7(2):151–174.

Biais, B., T. Foucault, and S. Moinas. 2015. “Equilibrium Fast Trading.” *Journal of Financial Economics* 116(2):292–313.

Bloomfield, R., M. O’Hara, and G. Saar. 2005. “The Make or Take Decision in an Electronic Market: Evidence on the Evolution of Liquidity.” *Journal of Financial Economics* 75(1):165–199.

Boehmer, E., G. Saar, and L. Yu. 2005. “Lifting the Veil: an Analysis of Pre-Trade Transparency at the NYSE.” *Journal of Finance* 60:783–815.

Booth, G., J.C. Lin, T. Martikainen, and Y. Tse. 2002. “Trading and Pricing in Upstairs and Downstairs Markets.” *Review of Financial Studies* 15:1111–1135.

Cao, C., O. Hansch, and X. Wang. 2009. “The Information Content of an Open Limit-Order Book.” *The Journal of Futures Markets* 29(1):16–41.

Chen, H., and P.M.S. Choi. 2012. “Does Information Vault Niagara Falls? Cross-Listed Trading in New York and Toronto.” *Journal of Empirical Finance* 19(2):175–199.

Chen, H., P.M.S. Choi, and Y. Hong. 2013. “How Smooth is Price Discovery? Evidence from Cross-Listed Stock Trading.” *Journal of International Money and Finance* 32:668–699.

Chen, W.P., and H. Chung. 2012. “Has the Introduction of S&P 500 ETF Options Led to Improvements in Price Discovery of SPDRs?” *Journal of Futures Markets* 32(7):683–711.

Chen, Y.L., and Y.F. Gau. 2010. “News Announcements and Price Discovery in Foreign Exchange Spot and Futures Markets.” *Journal of Banking & Finance* 34(7):1628–1636.

Chu, Q.C., W.G. Hsieh, and Y. Tse. 1999. "Price Discovery on the S&P 500 Index Markets: An Analysis of Spot Index, Index Futures, and SPDRs." *International Review of Financial Analysis* 8(1):21–34.

Cont, R., A. Kukanov, and S. Stoikov. 2014. "The Price Impact of Order Book Events." *Journal of Financial Econometrics* 12:47–88.

DeJong, F. 2002. "Measures of Contributions to Price Discovery: A Comparison." *Journal of Financial Markets* 5:323–328.

Eisler, Z., J.P. Bouchaud, and J. Kockelkoren. 2012. "The Price Impact of Order Book Events: Market Orders, Limit Orders and Cancellations." *Quantitative Finance* 12(9):1395–1419.

Engle, R.F., and C.W.J. Granger. 1987. "Co-Integration and Error Correction: Representation, Estimation, and Testing." *Econometrica* 55(2):251–276.

Engle, R.F., and J.R. Russell. 1998. "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data." *Econometrica* 66(5):1127–1162.

Frank, J., and P. Garcia. 2011. "Bid-Ask Spreads, Volume, and Volatility: Evidence from Livestock Markets." *American Journal of Agricultural Economics* 93:209–225.

Fricke, C., and L. Menkhoff. 2011. "Does the 'Bund' Dominate Price Discovery in Euro Bond Futures? Examining Information Shares." *Journal of Banking and Finance* 35(5):1057–1072.

Frijns, B., A. Gilbert, and A. Tourani-Rad. 2010. "The Dynamics of Price Discovery for Cross-Listed Shares: Evidence from Australia and New Zealand." *Journal of Banking & Finance* 34:498–508.

Frino, A., A. Lepone, and G. Wearin. 2008. "Intraday Behavior of Market Depth in a Competitive Dealer Market: A Note." *Journal of Futures Markets* 28:294–307.

Glosten, L. 1994. "Is the Electronic Open Limit Order Book Inevitable?" *The Journal of Finance* 49(4):1127–1161.

Gonzalo, J., and C. Granger. 1995. "Estimation of Common Long-Memory Components in Cointegrated Systems." *Journal of Business & Economic Statistics* 13(1):27–35.

Gonzalo, J., and J.Y. Pitarakis. 1998. "Specification via Model Selection in Vector Error Correction Models." *Economics Letters* 60:321–328.

Gould, M.D., M.A. Porter, S. Williams, M. McDonald, D.J. Fenn, and S.D. Howison. 2013 "Limit Order Books." *Quantitative Finance* 13(11):1709–1742.

Grammig, J., and F.J. Peter. 2013. "Telltale Tails: A New Approach to Estimating Unique Market Information Shares." *Journal of Financial and Quantitative Analysis* 48(2):459–488.

Harris, F.H., T.H. McInish, and R.A. Wood. 2002. "Security Price Adjustment across Exchanges: an Investigation of Common Factor Components for Dow Stocks." *Journal of Financial Markets* 5(3):277–308.

Harris, L., and V. Panchapagesan. 2005. "The Information Content of the Limit Order Book: Evidence from NYSE Specialist Trading Decisions." *Journal of Financial Markets* 8(1):25–67.

Hasbrouck, J. 1995. "One Security, Many Markets: Determining the Contributions to Price Discovery." *Journal of Finance* 50(4):1175–1199.

Hasbrouck, J. 2000. "Stalking the 'Efficient Price' in Market Microstructure Specifications: An Overview." Working paper, New York University.

Hasbrouck, J. 2004. "Liquidity in the Futures Pits: Inferring Market Dynamics from Incomplete Data." *The Journal of Financial and Quantitative Analysis* 39(2):305–326.

Hautsch, N., and R. Huang. 2012. “On the Dark Side of the Market: Identifying and Analyzing Hidden Order Placements.” Discussion Paper 2012-14, CRC 649, Humboldt-Universität zu Berlin.

Huang, R. 2002. “The Quality of ECN and NASDAQ Market Maker Quotes.” *Journal of Finance* 57:1285–1319.

Huang, R. 2000. “Price Discovery by ECNs and NASDAQ Market Makers.” Working Paper, Vanderbilt University, Nashville, TN.

Irwin, S.H., and D.R. Sanders. 2012. “Financialization and Structural Change in Commodity Futures Markets.” *Journal of Agricultural and Applied Economics* 44(3):371–396.

Kaniel, R., and H. Liu. 2006. “So What Orders Do Informed Traders Use?” *Journal of Business* 79(4):1867–1913.

Korczak, P., and K. Phylaktis. 2010. “Related Securities and Price Discovery: Evidence from NYSE-Listed Non-U.S. Stocks.” *Journal of Empirical Finance* 17(4):566–584.

Lehecka, G.V., X. Wang, and P. Garcia. 2014. “Gone in Ten Minutes: Intraday Evidence of Announcement Effects in the Electronic Corn Futures Market.” *Applied Economic Perspectives & Policy* 36(3):504–526.

Lien, D., and K. Shrestha. 2009. “A New Information Share Measure.” *Journal of Futures Markets* 29:377–395.

Lien, D., and K. Shrestha. 2014. “Price Discovery in Interrelated Markets.” *Journal of Futures Markets* 34(3):203–219.

Lien, D., and Z. Wang. 2016. “Estimation of Market Information Shares: A Comparison.” *Journal of Futures Markets* 36(11):1108–1124.

Liu, Q., and Y. An. 2011. “Information Transmission in Informationally Linked Markets: Evidence from US and Chinese Commodity Futures Markets.” *Journal of International Money and Finance* 30(5):778–795.

MacKinnon, J.G. 1994. “Approximate Asymptotic Distribution Functions for Unit-root and Cointegration Tests.” *Journal of Business and Economic Statistics* 12:167–176.

Madhavan, A., D. Potter, and D. Weaver. 2005. “Should Securities Markets be Transparent?” *Journal of Financial Markets* 8:265–287.

Martinez, V.H., and I. Roşu, 2011. “High Frequency Traders, News, and Volatility.” Working paper, Baruch College and HEC Paris.

Narayan, S., and R. Smyth. 2015. “The Financial Econometrics of Price Discovery and Predictability.” *International Review of Financial Analysis* 42:380–393.

Pascual, R., and D. Veredas. 2009. “What Pieces of Limit Order Book Information Matter in Explaining Order Choice by Patient and Impatient Traders?” *Quantitative Finance* 9(5):527–545.

Putnins, T.J. 2013. “What Do Price Discovery Metrics Really Measure?” *Journal of Empirical Finance* 23:68–83.

Rittler, D. 2012. “Price Discovery and Volatility Spillovers in the European Union Emissions Trading Scheme: A High-Frequency Analysis.” *Journal of Banking & Finance* 36(3):774–785.

Rock, K. 1996. “The Specialist’s Order Book and Price Anomalies.” Unpublished working paper, Harvard University, Boston, MA.

Seppi, D. 1997. “Liquidity Provision with Limit Orders and a Strategic Specialist.” *Review of Financial Studies* 10(1):103–150.

Stock, J.H., and M.W. Watson. 1988. "Testing for Common Trends." *Journal of the American Statistical Association* 83(404):1097–1107.

Tse, Y. 1999. "Round-the-Clock Market Efficiency and Home Bias: Evidence from the International Japanese Government Bond Futures Markets." *Journal of Banking and Finance* 23:1831–1860.

Vo, M.T. 2007. "Limit Orders and the Intraday Behavior of Market Liquidity: Evidence from the Toronto Stock Exchange." *Global Finance Journal* 17:379–396.

Wang, X., P. Garcia, and S.H. Irwin. 2014. "The Behavior of Bid-Ask Spreads in the Electronically-Traded Corn Futures." *American Journal of Agricultural Economics* 96:557–77.

Wuyts, G. 2009. "The Impact of Liquidity Shocks through the Limit Order Book." Working Paper. Department of Accountancy, Finance and Insurance, Faculty of Business and Economics, Katholieke Universiteit Leuven.

Yan, B., and E. Zivot. 2010. "A Structural Analysis of Price Discovery Measures." *Journal of Financial Markets* 13:1–19.