

Nearby and deferred quotes: What they tell us about linkages and adjustments to information

Mindy Mallory, Philip Garcia, and Teresa Serra

February 23, 2016

Abstract

Recently, the ‘Financialization’ of commodity futures markets, biofuel production, climate change, and rising demand potentially have imposed profound shifts in the way commodity futures markets operate. This article examines the commodity markets quote-by-quote to develop metrics on liquidity and transmission of information. The metrics are based on insights we combined from sequential trading models on single securities, index futures on a basket of securities, and special features of commodity futures markets. Correlation between quote revisions in nearby and deferred contracts measure information-based activity, and correlations between revisions of the time lagged nearby and deferred maturity measure the speed at which information is transmitted among the different futures maturities. Information based trading results in near perfect correlation between revisions to bids and offers in nearby and deferred contracts. Within one second information has been fully transmitted from nearby to deferred contracts.

Keywords: market, microstructure, futures, bid-ask-spread, corn

In italics are notes to ‘us’ about things that need to be added or clarified.

1 Introduction

There has been recent concern about whether and how the ‘Financialization of Commodity Markets’ has impacted market efficiency and efficacy in the traditional roles of risk mitigation, coordinating production, and coordinating consumption through time (Irwin and Sanders 2011; Cheng and Xiong 2013; Irwin and Sanders 2012; Henderson, Pearson and Wang 2015). Further, the recent increase in the production of biofuel from food commodities and volatile crude oil prices has changed the relationship between food and energy commodities (Serra and Zilberman 2013; Mallory, Irwin and Hayes 2012; Gardebroek and Hernandez 2013; Vacha et al. 2013; Avalos 2014; Trujillo-Barrera et al. 2012). Additionally, climate change, rising demand for agricultural commodities, and volatile inventories and exchange rates have imposed structural changes in commodity markets (Balcombe, Prakash and others 2011; Gilbert and Morgan 2010; Prakash, Gilbert and others 2011).

These issues represent potentially profound shifts in the way commodity markets operate, and the articles cited above have considered their implications. However, how these changes affect commodity markets tick-by-tick and quote-by-quote needs to be considered. Since global price discovery occurs on global futures exchanges for the major food commodities, a deep consideration of these changes on trading activity, patterns, and consequences is warranted. We use “high frequency data” (time stamped to the second), in order to allow for faster price change adjustments taking place after significant technical developments in trading platforms in the second half of the 2000s, characterized by high speed trading.

Price analysis can be classified into structural and non-structural studies. While structural models rely on economic theory, non-structural analyses identify empirical regularities in the data. The approach throughout this article is non-structural. We employ this approach primarily because there is scant market microstructure literature that was developed with the particular characteristics of commodity futures markets in mind. In this article, we are motivated to develop tools with which we can consider the microstructure implications of important trends like the financialization of commodity futures or links between energy and food commodities. Our purpose is to develop initial metrics of information- versus liquidity-based activity in commodity markets. We anticipate this

work will lead to future developments in the microstructure of commodity markets literature.

Even how to develop simple metrics of information- versus liquidity-based activity from standard microstructure models is not obvious because standard models of trading securities are not necessarily directly applicable to commodity futures markets. For example, in commodities futures markets several contracts with different maturities trade in the marketplace, each reacting to information and liquidity-based shocks. Each contract reacts differently to information-based shocks because there is a cost to store the physical commodity through time. Therefore, an information based shock that is most pertinent to the nearby contract may have a large effect on the nearby futures contract, but a dampened effect in more distant maturing contracts. For example, information about shortages and surpluses in the near-term are mitigated in distant contracts by high (low) prices today that will ration (encourage) consumption and therefore mitigate price effects. Further, each contract maturity attracts different levels of liquidity. Generally liquidity decreases at more distant contracts except for commodities produced once per year. Corn, for example is harvested in the fall and the ‘new crop December contract’ routinely attracts more volume than the May or July contract even when the May or July contract expires sooner.

The metrics we develop in this article on liquidity and transmission of information are based on insights we combined from the sequential trading models on single securities, index futures based on a basket of securities, and some of the features of commodity futures markets described in the preceding paragraph. Using the standard sequential trading result that quote revisions only occur if liquidity providers have updated their beliefs about the value of the security after observing order flows, the correlation between quote revisions in nearby and deferred contracts can be used to measure information-based activity, and correlations between revisions of the time lagged nearby and deferred maturity can be used to measure the speed at which information is transmitted among the different futures maturities. This metric is sensible in commodity futures markets but not in a market for a single security because futures markets have multiple maturity contracts that should respond to information in a very similar and predictable way.

We find within one second information has been fully transmitted from nearby to deferred contract maturities. Without a model explicitly relating these correlations to a sequential trading model, we cannot interpret much about the size of the correlations or the speed of information transfer, but these facts certainly shed light on how the markets behave recently. Future research is

needed to develop a sequential trading model for commodity futures markets where more insights about determinants of behavior can be gleaned.

The remainder of the article is organized as follows. First we provide a background of the sequential trading and index futures microstructure literature and describe the conceptual framework that motivates our interpretation of correlations of quote revisions as a metric of information-based activity. Next we describe the data and report the results of our analysis. Finally, we make some concluding remarks.

2 Literature Review

The literature on how information affects liquidity in securities markets is long and rich.¹ Bagehot (1971) is regarded as the first to demonstrate that a bid-ask spread (BAS) arises when asymmetric information is present even if inventory and transactions costs are assumed to be zero. Copeland and Galai (1983) build upon Bagehot's work by assuming that a specific proportion of traders were informed. Knowing this the market maker adjusts his quoted bids and offers to maximize expected profit. Copeland and Galai's model, however, does not account for the fact that the trades themselves can reveal information about whether or not traders are informed. Glosten and Milgrom (1985) formalize this concept and develop a model where the market maker adjusts his beliefs based on the trades that occur. The market maker knows that at least some of the traders are informed so sell orders revise the market maker's belief downward about the value of the security and buy orders revise his belief upward. They show that the spread is increasing in the proportion of informed traders, and there is a point at which too many informed traders require the market maker to set the spread so wide that trade does not occur and the market halts (an example of the famous "Market for Lemons" described by Akerlof (1970)).

Easley and O'Hara (1987) and Easley and O'Hara (1992) incorporate trade size and its effect to a model similar to Glosten and Milgrom. A market maker must set breakeven bid and offer quotes knowing that he faces a certain proportion of informed traders who only trade if they receive a signal that an information event has occurred and a certain proportion of uninformed traders who do not receive an information signal but occasionally need to trade for liquidity reasons. Both informed

¹The interested reader can refer to O'Hara (1995) for an excellent and detailed overview of the evolution of this literature.

and uninformed traders can choose between a large and small block trading size. This model setup leads to two types of equilibria: a separating equilibrium where informed traders only trade in large quantities and a pooling equilibria where informed traders may trade both large and small quantities. This model setup of information uncertainty and asymmetric information leads to the market maker updating his beliefs about the value of the security (and therefore his quotes) based on the order flow he observes in the market. For example, in a separating equilibrium a large trading block causes the market maker to revise upward his expectation that an information event has occurred (since informed traders do not transact at small sizes). This results in a tradeoff for the informed trader. Transact at ‘good’ price with one large lot, or transact with several small lots at ‘good’, but worsening, prices as the market maker learns slowly from the informed traders’ one-sided order flow.

Hasbrouck (2006) provides an overview of how Easley, Hvidkjaer and O’Hara (2002) and Easley, Kiefer and O’Hara (1997) use the Easley and O’Hara models of informed trading to develop a measure of the probability of informed trading (PIN). This measure, though, is estimated solely based on the sequence of order arrivals, where a trade is labeled as buyer initiated if the trade occurs above the midpoint of the quoted spread and seller initiated if the trade occurs below the midpoint of the quoted spread. Numerous studies have documented that there may be problems with downward bias in the estimated PIN (Yan and Zhang 2012; Vega 2006; Boehmer, Grammig and Theissen 2007) and estimating information based trading in this way ignores some aspects of futures markets discussed above that are not present in securities markets. For these reasons we seek an alternative to the PIN measure of information based trading in commodity futures.

To our knowledge there are no market microstructure models that explicitly take into account the features of commodity futures markets. The closest models come from work on index futures that cover a basket of securities. Most prominent is the work by Kumar and Seppi (1994) who assume N different non-dividend paying securities and an index futures contract on a buy-and-hold portfolio of a subset of these stocks. In Kumar and Seppi’s model, specialists in the cash market observe a signal, and floor traders of the futures index observe a signal about the value of the index but not the individual securities. A key feature they build into the model is a lag in the information transmittal between the cash and futures markets because specialists only observe order flows from their own market, and not the other. This lag in information transmittal allows for arbitrageurs, who possess faster telecommunication technologies, to learn from transactions in both markets and

make profitable trades in the cash and futures markets. In common futures market parlance, these ‘arbitrageurs’ are analogous to spread traders who trade in both nearby and deferred contracts hoping to profit on relative price movements.

There are some important distinctions between the arbitrageurs as proposed in the Kumar and Seppi model and spread traders in a futures market. Namely, the basis between a composite of cash security prices and the price of a futures index of the same basket should behave in very predictable ways (the basis, in theory, should only vary with interest rates and expected changes in dividend yields if information is symmetric). In contrast, the spread between the prices of two commodity futures contracts with different maturities depends on many more uncertain structural variables: e.g., domestic and international consumption, exchange rates, production or distribution bottlenecks, and weather. The arbitrageurs in Kumar and Seppi’s model need only to wait for others in the marketplace to learn to profit. The futures market spread trader entertains much more risk in betting on relative price changes between two futures maturities.

However, Kumar and Seppi’s model yields two insightful empirical predictions that provide insight. Their Proposition 1 states that it may not be optimal for the arbitrageur to hold a perfectly offsetting position. Because there is a time lag over which the specialists learn about realized market signals in the other market, it may be optimal for the arbitrageur to hold a net long or net short position in the cash or index futures market. As noted above, the commodity futures market is even noisier and so this proposition may hold for commodity futures as well. Next, their Proposition 2 states that the correlation between arbitrageurs’ stock and futures orders is negative, meaning if they are long the in the cash basket of securities they are likely to be short the basket of index futures, or vice-versa. If arbitrageurs only made perfectly offsetting orders in the cash and futures markets, this would trivially be a perfect negative correlation. What this behavior means for bids, offers, and relative transaction prices in the cash and futures market is unclear because the arbitrageurs’ activities can cause the cash and futures prices to come together or cause them to diverge.

In the next section we draw insights from the sequential trading models described above to generate empirical predictions about the correlations between revisions to bids and offers of nearby and deferred maturity commodity futures contracts.

3 Conceptual Framework

In this section we develop a conceptual framework for how the role of liquidity-based activity versus information-based activity should affect quote revisions in a commodity futures market. Using insights from the Easley and O'Hara and Kumar and Seppi models, along with features of commodity futures markets, we generate empirical predictions about the correlations between revisions to quotes in the nearby and deferred maturity commodity futures contracts.

First, consider an absence of information. In the Easley and O'Hara sequential trader models, the market maker revises his quotes only when he updates his belief that the value of the security has changed. Therefore, we interpret no changes in revisions to bids (offers) as indicative of no information having arrived to the market. Any transactions that occur at these prices, the market maker believes are conducted by uninformed traders demanding liquidity.

Conversely, when we observe revisions to the bid or offer, we can infer that the market maker from the Easley and O'Hara models has updated his beliefs about information arrival to the market based on past order flows. These revisions to the bid and offer we interpret as indicative of information having arrived to the market.

Now we discuss features of futures markets that we can utilize when considering revisions to nearby and deferred contract quotes. First, in actively traded commodity futures markets there is no market maker, but there are entities who actively supply liquidity to the market under a variety of motives. Since the Easley and O'Hara models consider a competitive market maker, it is irrelevant whether there is one market maker in the traditional sense or a large number of liquidity providers 'making markets'. Second, when market makers revise their beliefs that an information event has arrived to the market, they know it affects futures contracts of all maturities so quotes must be revised in all contracts.

This should induce a high degree of correlation between bid and offer revisions when an information event arrives. Further, one market maker would revise bids and offers on futures contracts of all maturities at the same time they update beliefs about an information event having occurred. As a practical matter with many entities making markets in the futures contracts, it is not clear that this updating will happen in all maturities simultaneously. Therefore, it is of interest to measure correlation contemporaneously and at increasing time lags of the nearby to determine how

quickly beliefs are updated in deferred maturities relative to the nearby.

4 Data

The data used in this analysis come from the CME Group’s Top of Book (BBO) database for corn futures contracts from 01/14/2008-11/4/2011. The data contain the best bid, bid volume, best offer, offer volume, last trade price, and last trade volume of the order book for each active futures contract, time-stamped to the second. For those familiar with the CME Group’s Time & Sales database, the BBO contains the Time & Sales data as well as information about the best bids and offers on the top of the limit order book.

The time period examined was characterized by volatility, uncertainty, and rapid increases in prices in the beginning and end of the sample. Prices increased to a peak of approximately \$7.50 per bushel in 2008, a time a time which saw a broad class of commodity markets exhibiting similar rapid price increases. Then a relatively stable period from 2009 and 2010 saw prices within a relatively tight range of \$3.00 to \$4.50 per bushel. In the final year of the sample, uncertainty and rapid price increases reigned again as worries about a smaller than anticipated crop and small ending stocks drove prices to nearly \$8.00 per bushel. While prices were volatile, the share of contracts traded on the CME’s electronic trading platform, Globex, had already stabilized to nearly 90% by 2008 (Peterson 2015). So any effects we study should not be related trading infrastructure changes that may have occurred during the migration of volume to the electronic exchange.

The data are time-stamped to the second, but trades and updates to the top of the book routinely occur more rapidly than once per second. This results in several updates to the top of the book displaying the same time stamp. This requires us to either aggregate to the second, or to simulate sub-second time stamps (Hasbrouck 2015; Wang 2014). Since we calculate correlations between updates to the top-of-the book for several contract maturities, simulation would need to preserve (at least) the order in which updates arrived to each respective contract to be useful. Since preserving the order in which updates to the top of the book is impossible, we aggregate to the second.²

Further, we exclude days on which there was a limit price move in any of the contracts, since when prices are locked at the limit, calculating correlations is impossible (Dates deleted due to limit

²We take the last entry on each timestamp for the aggregation.

price moves and the corresponding information events, if known, are as follows: 1/12/2010, revision to a Crop Production report; 3/31/2011, Prospective Plantings report; 6/30/2011, Planted Acres report; 10/8/2010, World Agricultural Supply and Demand Estimates (WASDE); and 12/9/2010, WASDE). Also, we exclude 4/5/2010, because there was an unusually high number of revisions to the best bid and best offer. Since we were not able to process all of the data for this day in a reasonable amount of computing time, we drop this day from our sample. Additionally, 7/5/2011 was an unusually light trading day after the Fourth of July holiday and resulted in no data for the third to mature contract, so we dropped this day as well.

Table 1 shows the first ten entries to our data after manipulating the raw BBO data set from CME Group to display the entire top of the book on one line with the appropriate time stamp. The first column is the time-stamp, the second column is the trade sequence number, which the CME Group gives to individual trades to identify separate orders that arrive on the same second. The third column, SYMBOL, identifies which futures maturity the observation represents. In this case, 1003 stands for March 2010, with the first two characters representing the year and the second two characters representing the month. The fourth column, OFRSIZ, is the number of contracts quoted at the best offer price. The fifth column, OFR, is the best offered price. The sixth column, BIDSIZ, is the number of contracts quoted at the best bid price; the last column, BID, is the best bid price.

For each date in our sample, we define the first to mature (nearby), one, two, and three contracts deferred. To avoid undesirable delivery month effects, we roll these definitions at the end of the month prior to contract maturity (For example, on March 1st the May futures contract is considered the nearby as opposed to the March futures contract, and so forth).

Describe new figure 1

5 Analysis

Our analysis considers the correlation of revisions to nearby contract bids (offers) with one, two, and three deferred contract bids (offers), respectively. We described in the Conceptual Framework section that when information arrives to the market, it usually will affect the entire forward curve in the same direction. In other words, information that raises the best bid (offer) in the nearby contract, should usually raise the best bid (offer) in the deferred contracts as well. Linkages between the nearby

and deferred contracts can be measured with simple correlations without making the distributional assumptions required by a more intricate econometric model. One of the key methodological issues when it comes to assessing high frequency data is the non-normality of price data that complicates proper modeling and requires the use of appropriate methodologies for appropriate analysis (Aït-Sahalia, Mykland and Zhang 2005; Andersen et al. 2001; Easley, Prado and O'Hara 2012; Hasbrouck 2013; Lee and Mykland 2008; Lehecka, Wang and Garcia 2014).

We have two primary objectives: 1) calculate the strength of correlations between the order books of the nearby and deferred contracts, and 2) measure (or bound) the time it takes for information to be transmitted from nearby to deferred contracts. To measure the first, we calculate contemporaneous correlations between the log changes of the bid (offer) in the nearby and the deferred contracts. Then, to measure the second we calculate the correlation between time lagged log changes of the bid (offer) of the nearby with log changes of the bid of the deferred contracts. We lag the nearby by one second and ten seconds. The time lagged correlations provide a measure of how long it takes for information to be transmitted from nearby to the deferred contracts. The logic is that if we observe contemporaneous correlation between the nearby and deferred contracts, we can search for the time lag at which we observe the correlation disappearing. We conclude that information has been fully transmitted when the time lagged nearby and deferred contract order book revisions become uncorrelated. Conversely, we may observe that there is no contemporaneous correlation, but there is lagged correlation.

Since the corn futures contract experiences non-uniform trading volume throughout the day, there may be time of day effects in the strength and rate at which information is transmitted through the futures market. To measure how the correlations change throughout the trading day, we divide the day into ten minute intervals (bins) and calculate correlations between the nearby and the deferred contract bids (offers) for each interval. Ten minutes was shown to be long enough for market adjustment to take place in Lehecka (2014). This allows us to detect if there are any discernible patterns to the correlations over the trading day. Further, since we calculate one correlation per bin per day we also recover a whole distribution of the correlations for each time interval.

5.1 Definition of Information-Based Activity

As mentioned, it is common to have multiple revisions to the order book on the same second (and consequently receive the same time-stamp in the data). The converse is also true, however. It is also common for a number of seconds to transpire before the top of the order book is revised. This results in our variables, log changes in best bid (offer), containing many zeros. How these zeros are distributed between the contracts is related to the concepts of liquidity-based versus information-based activity discussed in the conceptual framework.

To fix ideas consider the case where we intend to calculate the contemporaneous correlation between the nearby and the deferred contracts. There are three possibilities for the presence of zeros. On any time stamp one of the three situations may occur: 1) neither the nearby nor the deferred has a zero log change in the bid (offer), 2) either the nearby or the deferred has a zero log change in the bid (offer), but not both, or 3) both the nearby and the deferred have a zero log change in the bid (offer).

Based on the definition of liquidity-based activity and information-based activity in the conceptual framework, we present a case for interpreting (1) as information-based activity, (2) liquidity-based activity, and (3) liquidity-based activity.³

The intuition is that if both the nearby and deferred contracts experience a revision in the same direction, they are likely responding to the arrival of information to the marketplace, and best bids (offers) adjust accordingly. This is in contrast to the case where one of the two contracts experiences a revision and the other contract has no change. If one contract experiences a revision in the best bid (offer) and the other does not, it is likely that the revision results from a liquidity-based order in an effort to exit their position.

If this intuition is correct, it is informative to consider only time-stamps for which both contracts experienced a revision - that is isolating what we are referring to as information-based activity. In the next section we will present results based on all data in our sample and the sub-sample where only the time-stamps for which both contracts experience a non-zero revision to the top of the book are included as separate cases.

³In the third case, no activity at all is observed in the quoted price changes, but quoted quantities may have changed due to new limit orders arriving, limit orders being cancelled, or market orders arriving taking some of the quoted quantities off the book. This is indicative of liquidity-based activity as well.

5.2 USDA Announcement Days

On USDA report announcement days there is often a significant amount of information that market participants receive at the same time, causing large price fluctuations and larger than usual trading volumes. Therefore, in our analysis we also separate out days on which major USDA reports are released and calculate the same correlations described above. During our sample period, the USDA reports we include were released at 8:30 am CST, before the day trading session began.

5.3 Considering the Correlation of Nearby Bid with Deferred Offer and Nearby Offer with Deferred Bid

Nearby offer and deferred bid bid combine to form the bid for the calendar spread, and nearby bid and deferred offer combine to form the offer for the calendar spread. Discuss how for this reason we consider the bid-to-offer and offer-to-bid correlations.

List correlations calculated here and description of expectations

6 Results

In this section we present a series of charts that display the calculations described in the preceding section. For each set of figures, panel A is produced with correlations between log changes to the bid in the nearby and log changes to the bid in the deferred. Panel B is produced in the same way as panel A, but offers are used. Correlations are calculated over ten minute intervals starting at 9:30am Central Standard Time, the beginning of the daytime trading session for CBOT corn futures. One correlation is calculated per day per ten minute bin, so for every ten minute bin we recover a distribution of correlations. The dots represent the mean of the distribution and the bars represent one standard deviation.

6.1 Contemporaneous Correlation - Information-Based Activity Only

In figure 2 contemporaneous correlation between the nearby and one, two, and three deferred maturity contracts are displayed. Calculations are made based on time-stamps where both the nearby and

deferred maturity experience non-zero revisions to the best bid (top panel) or offer (bottom panel). The contemporaneous correlations between each nearby and deferred contract pairs are very close to one for both best bids (top panel) and best offers (bottom panel). The exception being that there is a slight dip in correlations at the beginning and end of the trading day.

This implies that in the event both contracts experience revisions to their respective limit order books, they are revised in lockstep. While some of this correlation is artificially induced due to the tick structure of price changes in this market (prices move in 0.25 cent increments.), not all of it can be attributed to that. Additionally, since our data is only time-stamped to the second, we may be missing nuance that would be captured with data time stamped to the millisecond. Regardless, the result is surprisingly strong and supports interpreting this activity as information based.

It is interesting that the distribution of correlations between the nearby and deferred contract bids are at such similar levels, hovering very close to one. Also, the correlation between nearby and two contracts deferred is highest, nearby and three contracts deferred is next highest, and the nearby and one contract deferred is lowest, across all ten minute intervals. Since initially this ordering seems puzzling, we ran a Welch’s two-sample t -test with unequal variances to see if the means were significantly different from one another. The p-values from these results are shown in table 2. Comparing the mean of the correlations of the nearby and one deferred with the correlations of the nearby and two deferred, we reject the hypothesis of equal means in most of the time bins. In columns two and three we see less evidence of unequal means toward the midday and close of trading, but there is still evidence of a difference in means (rejecting the null hypothesis) for time bins earlier in the trading day. We are not sure what is causing this particular ordering of the means of the correlations. We expected either a random ranking among correlations with the different deferreds or that the highest correlations would result from the nearest to mature contracts.

We removed all but the ‘information based activity’ for brevity, but another question is if we remove the figures where we include all bid revisions (the old figure 1 where means are around 0.5 to 0.6 for example) do we lose the impressiveness that this current figure one using ‘information based activity’ only is so close to one?

6.2 Contemporaneous Correlation on USDA Report Days

Figure 2, is exactly analogous to figure 1 except that it focuses on USDA report days. We see a remarkably similar depiction compared to figure 2, in that the correlations hover near one throughout the trading day. If there had been a difference in the pattern of correlations on USDA report days, one would expect the first ten minutes of trading to display the largest effect. There is visibly more variation in the means of these distributions, resulting from the small sample of report days we average over, especially for time bins 09:30:00 and 09:50:00am. These have error bars that are much wider than the rest reflecting the influence of a small number of large moves in the opening minutes of trading for a few USDA report days.

We suspect two primary reasons that the full sample and USDA report day results are so similar: 1) Since we removed days where the report release corresponded to limit price moves, we systematically removed report days where the most important information was conferred on the market. It is possible that the remaining days corresponded to USDA reports that contained relatively less relevant market information, or they contained information market participants widely expected beforehand. 2) Since USDA reports were released prior to the market open during this time period, the information may have already been fully incorporated by market participants by the time the market opened, resulting in no discernible difference in the pattern of correlations in the first (and subsequent) time bins.

6.3 Contemporaneous and Time-lagged Correlations

An important contribution of this analysis is to provide insights on the speed at which information is transmitted from the nearby to the deferred contracts. To measure this, we lag the nearby series of log changes in the bid (offer) and calculate the correlation with the deferred bids (offers). This allows us to determine the length of time it takes for information to be fully transmitted to the deferred contracts. The assumption here is that the length of time it takes for the revisions to the top of the nearby limit order book to become uncorrelated with revisions to the top of the deferred limit order books is the length of time it takes for information to be transmitted between the two markets.

Figure 4 contains the correlations between the nearby and one deferred contracts. The graph shows the contemporaneous correlation from figure 2 as a reference, and correlations generated by

lagging the nearby by one second and ten seconds respectively. Here we expected to see a clear pattern of decreased correlation as we increased the length of the time lag in the nearby - reflecting that information is transmitted from nearby to deferred contracts over a number of seconds. However, we see that the correlation drops to zero with a lag of one second, which in this data set is the shortest time lag possible. We generated this figure for USDA report days only in figure 4. Similar to our result before, it is basically the same as figure 3 with more variation due to the small sample.

There are three possible explanations for this. First, it is possible that there is in fact a clear and decreasing correlation between lags of the nearby and the deferreds, but it can only be observed on mili- or micro-second time stamps. Then, when aggregating to the nearest second, we only observe contemporaneous correlation, but zero correlation even at the shortest possible time delay. Second, we have been explicitly assuming that price discovery happens in the nearby contract when we lagged the nearby contract instead of the deferred contract. If price discovery happens in the deferred contract, and takes time to fully to be incorporated into the nearby contract, then we would observe non-zero correlation between nearby quote revisions and lagged deferred quote revisions. This is presented in figure 5. Third, this would also occur if linkages between the nearby and deferred contracts were immediately enforced by spread traders. Surely the spread trade is an important component that keeps nearby and deferred contracts linked in economically meaningful ways. However, a spread trade is entered as a buy (sell) in the nearby and an sell (buy) in the deferred contract. Until now, we have been measuring correlations bib-to-bid and offer-to-offer between the nearby and deferred contracts. In figure 6, we will explore the nearby bid-deferred offer and nearby offer-deferred bid correlations to more directly measure the effect of spread traders in transmitting information up the forward curve.

Figure 5 displays the means and error bars of the calculated correlations between nearby bid (offer) revisions and lagged first deferred bid (offer) revisions. These results are nearly indistinguishable from the results in figure 3 with assumed price discovery happened in the nearby contracts. A very small mean positive correlation is detected when the first deferred is lagged by one second, but the size of the correlation does not appear to be meaningful larger than zero.

In figure 6 we generate a figure analogous to figure 1 in that it calculates the contemporaneous correlations between the nearby and deferred revisions to the top of the book. Figure 6, however, calculates the correlation between revisions to the nearby by bid with the deferred offer in the top

panel, and nearby offer with deferred bid in the bottom panel. Here, as in figure 1 we see correlations hovering near one. We observe that the correlations when the nearby is lagged by both one second and ten seconds are near zero, however, the mean of the one second lagged correlations are clearly higher than the ten second lags and larger than zero.

7 Conclusions

I have not updated conclusions to reflect any of the changes above.

Recent developments in commodity markets make it important to assess price adjustment patterns. We focused this paper on the corn market because it has experienced some of the most pronounced changes in recent years. We gleaned insights from the sequential trading market microstructure literature to generate metrics of informed versus liquidity trading in commodity futures markets. Sequential trading models allow liquidity providers to learn about the existence of information arrivals and their directional implications for security prices. From these models we infer that market makers detect no new market information if we observe no changes to the best bid or best offer in the limit order book. This is because in sequential trading models, the market maker learns about the probability of an information event from trader order flows and revises his breakeven bids and offers accordingly.

We use simple correlations between revisions to the best bid (offer) in the limit order book as our metric of information-based activity in the market. Our results for CBOT corn indicate that the mean contemporaneous correlation between the nearby and all deferred contracts considered was between 0.5 and 0.6 throughout the trading day. When we measured the correlation using information based activity only, we found the mean correlation between the nearby and all deferred contracts are nearly 1, indicating that when information arrives to the market, liquidity providers in contracts of all maturities revise their quotes in lockstep.

To measure the speed of information transfer from nearby to deferred maturities, we lagged the nearby by one and ten second respectively and computed correlations in revisions to the best bid (offer) with deferred contracts. We find that even at a one second lag, the shortest time lag possible with this data set, the correlation between revisions to the best bid and best offer dropped to zero. This indicates that by one second, any information that arrived to the market had been

fully transmitted across all contract maturities.

These results indicate that we can probably learn much from trades and quotes data about how many informed traders there are in a marketplace, but it is clear that much additional work in this area is needed. We borrowed from the sequential trading in securities literature, but future research is needed to develop sequential trading models specific to commodity futures markets. Such a model would provide richer insights into trading and behavior of markets.

8 Tables

Table 1: First ten entries in our data set.

| | ID | SYMBOL | OFRSIZ | OFR | BIDSIZ | BID |
|---------------------|-------|--------|--------|--------|--------|--------|
| 2010-01-04 09:30:00 | 98790 | 1003 | 1475 | 423.75 | 1188 | 423.75 |
| 2010-01-04 09:30:00 | 98800 | 1003 | 1483 | 423.75 | 1188 | 423.75 |
| 2010-01-04 09:30:00 | 98810 | 1003 | 1483 | 423.75 | 1197 | 423.75 |
| 2010-01-04 09:30:00 | 98820 | 1003 | 1486 | 423.75 | 1197 | 423.75 |
| 2010-01-04 09:30:00 | 98830 | 1003 | 1486 | 423.75 | 1231 | 423.75 |
| 2010-01-04 09:30:00 | 98840 | 1003 | 1494 | 423.75 | 1231 | 423.75 |
| 2010-01-04 09:30:00 | 98850 | 1003 | 1496 | 423.75 | 1231 | 423.75 |
| 2010-01-04 09:30:00 | 98860 | 1003 | 1510 | 423.75 | 1231 | 423.75 |
| 2010-01-04 09:30:00 | 98870 | 1003 | 1510 | 423.75 | 1233 | 423.75 |
| 2010-01-04 09:30:00 | 98880 | 1003 | 1520 | 423.75 | 1234 | 423.75 |

Notes: ID = CME's trade sequence number, Symbol = Contract expiration year (2010) and month (March), OFRSIZ = Number of contracts at the best offered price, OFR = Best price offered (cents per bushel), BIDSIZ = Number of contracts at the best bid price, BID = Best price bid (cents per bushel).

Table 2: P-Values from Welch two-sample t-tests on differences in correlation means displayed in figure 1

| | Near & 1 - Near & 2 | Near & 1 - Near & 3 | Near & 2 - Near & 3 |
|----------|---------------------|---------------------|---------------------|
| 09:40:00 | 0.06 | 0.00 | 0.00 |
| 09:50:00 | 0.00 | 0.07 | 0.00 |
| 10:00:00 | 0.00 | 0.00 | 0.00 |
| 10:10:00 | 0.00 | 0.01 | 0.00 |
| 10:20:00 | 0.00 | 0.00 | 0.00 |
| 10:30:00 | 0.00 | 0.00 | 0.00 |
| 10:40:00 | 0.00 | 0.00 | 0.04 |
| 10:50:00 | 0.00 | 0.00 | 0.04 |
| 11:00:00 | 0.00 | 0.00 | 0.13 |
| 11:10:00 | 0.00 | 0.00 | 0.37 |
| 11:20:00 | 0.00 | 0.00 | 0.09 |
| 11:30:00 | 0.00 | 0.01 | 0.28 |
| 11:40:00 | 0.00 | 0.02 | 0.31 |
| 11:50:00 | 0.00 | 0.06 | 0.21 |
| 12:00:00 | 0.00 | 0.11 | 0.13 |
| 12:10:00 | 0.01 | 0.09 | 0.37 |
| 12:20:00 | 0.01 | 0.07 | 0.35 |
| 12:30:00 | 0.01 | 0.17 | 0.22 |
| 12:40:00 | 0.02 | 0.16 | 0.36 |
| 12:50:00 | 0.03 | 0.18 | 0.38 |
| 13:00:00 | 0.02 | 0.23 | 0.28 |
| 13:10:00 | 0.10 | 0.54 | 0.31 |
| 13:20:00 | 0.73 | 0.06 | 0.13 |

Notes: H0: Equal means, H1: Unequal means. ‘Near & 1’ stands for the correlations between returns of the nearby quotes and returns of the first deferred quotes, ‘Near & 2’ stands for the correlations between returns of the nearby quotes and returns of the second deferred quotes, and ‘Near & 3’ stands for the correlations between returns of the nearby quotes and returns of the third deferred quotes. The columns labeled with the convention Near & 1 - Near & 2 contain p-values of a Welch two-sample t-test on differences in the means displayed in figure 2.

9 Figures

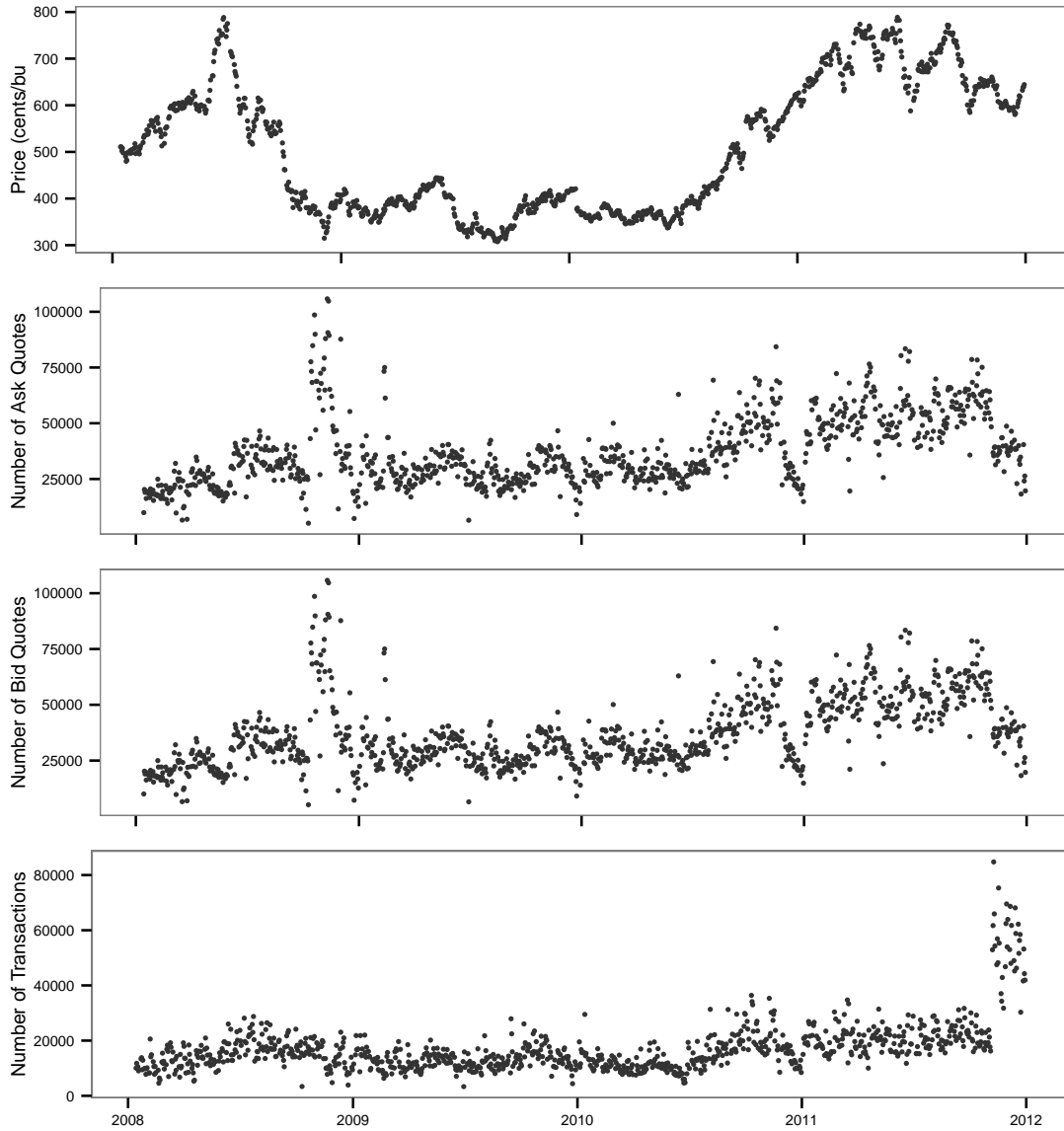


Figure 1: Price Levels, Number of Ask Quotes, Number of Bid Quotes, and Number of Transactions

Figure displays data for the corn futures market from 1/14/2008 to 12/31/2011 for the nearby contract. The September contract is excluded due to the possibility of ‘old crop’ and ‘new crop’ both being delivered on this contract. To form the continuous nearby series contracts are rolled to the next contract on the 20th of the month prior to the delivery month.

Figure numbers have not been updated to reflect the addition of the figure above. Nor has a discussion of the figure above been added to the manuscript yet.

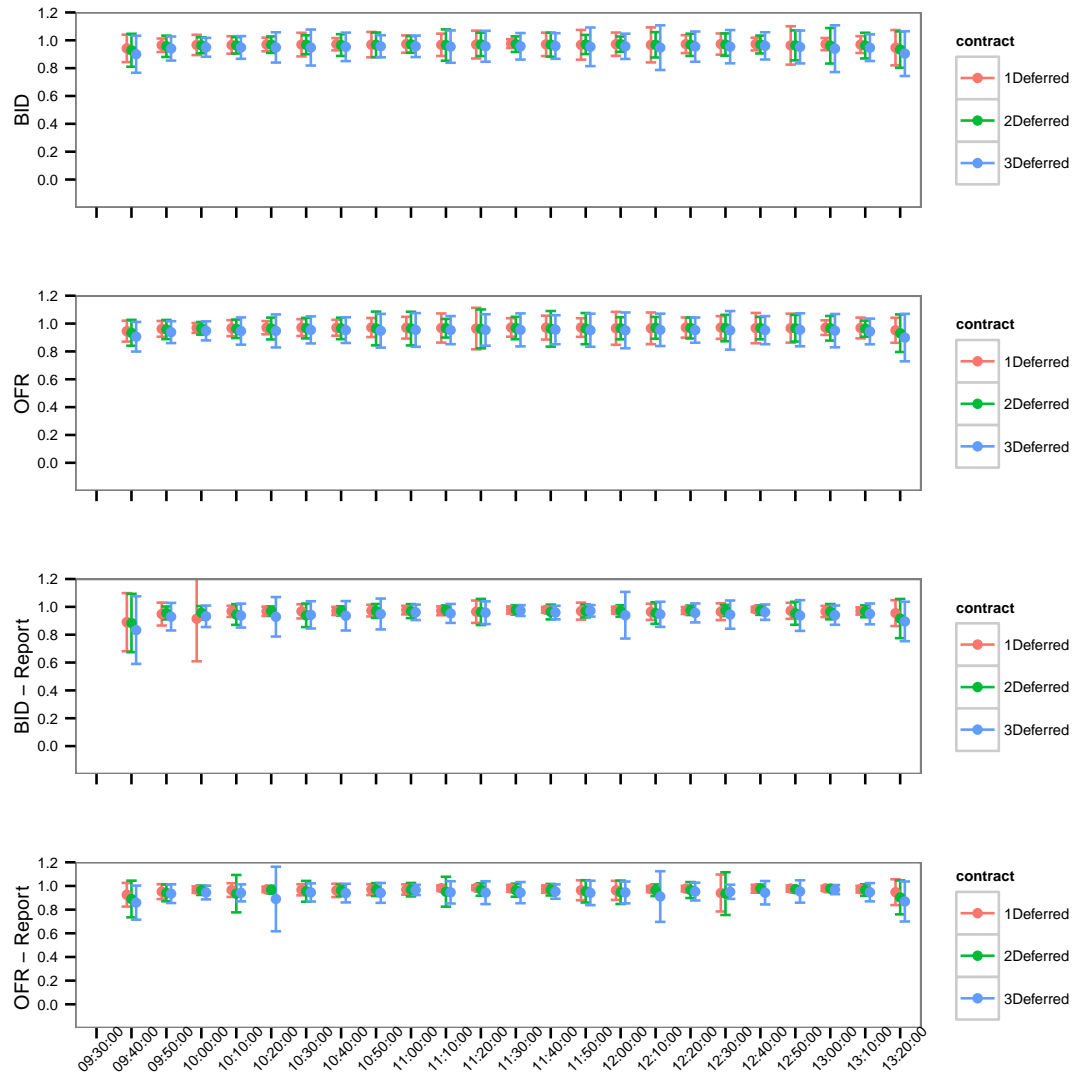


Figure 2: Contemporaneous Correlation between Revisions to Quotes in Nearby and Deferred Contracts

Mean correlations and one standard deviation error bars over all days in the top two plots. Mean correlations and one standard deviation error bars in the bottom two plots.

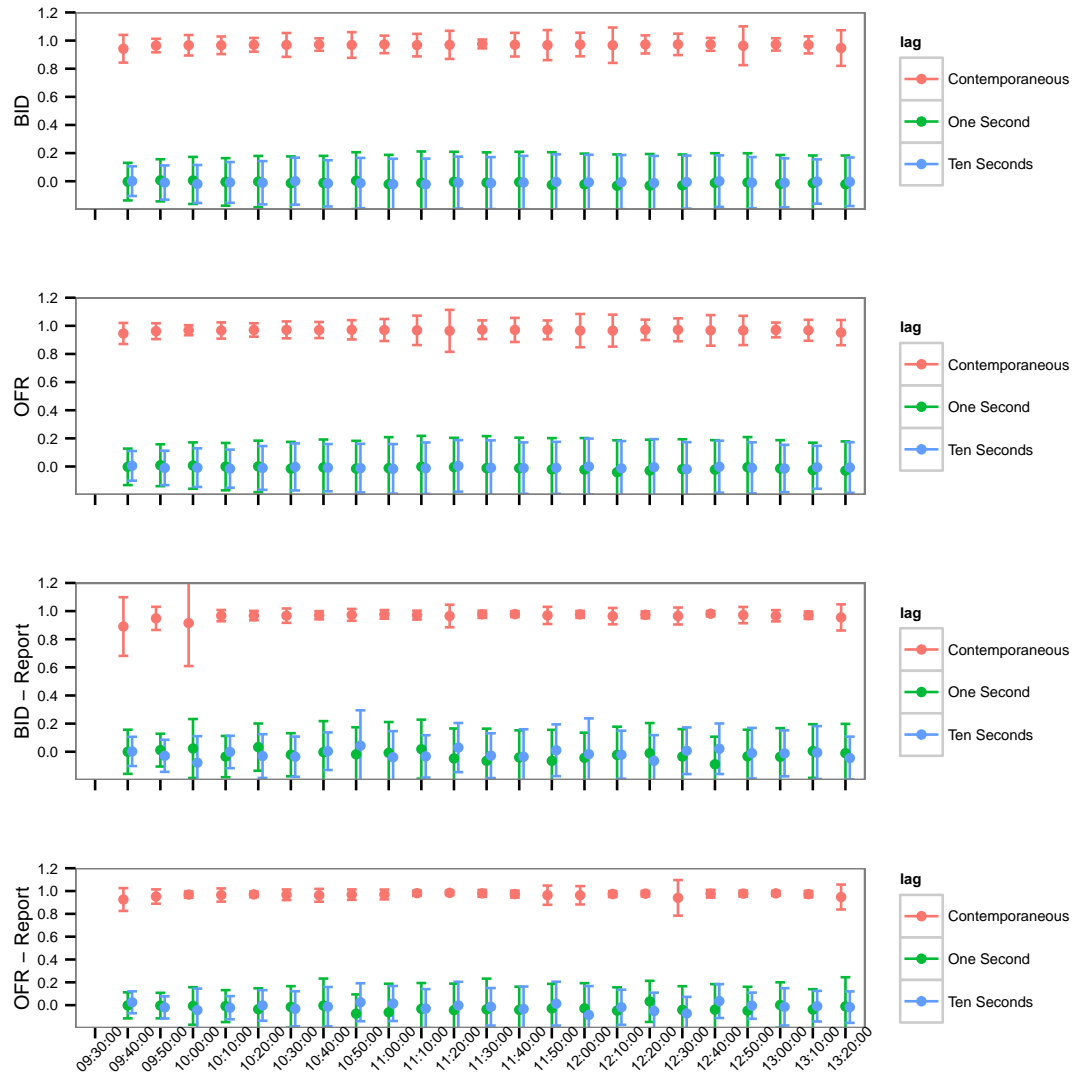


Figure 3: Time-Lagged Correlation between Revisions to Quotes in Nearby and 1 Deferred Contracts

Mean correlations and one standard deviation error bars over all days in the top two plots. Mean correlations and one standard deviation error bars in the bottom two plots.

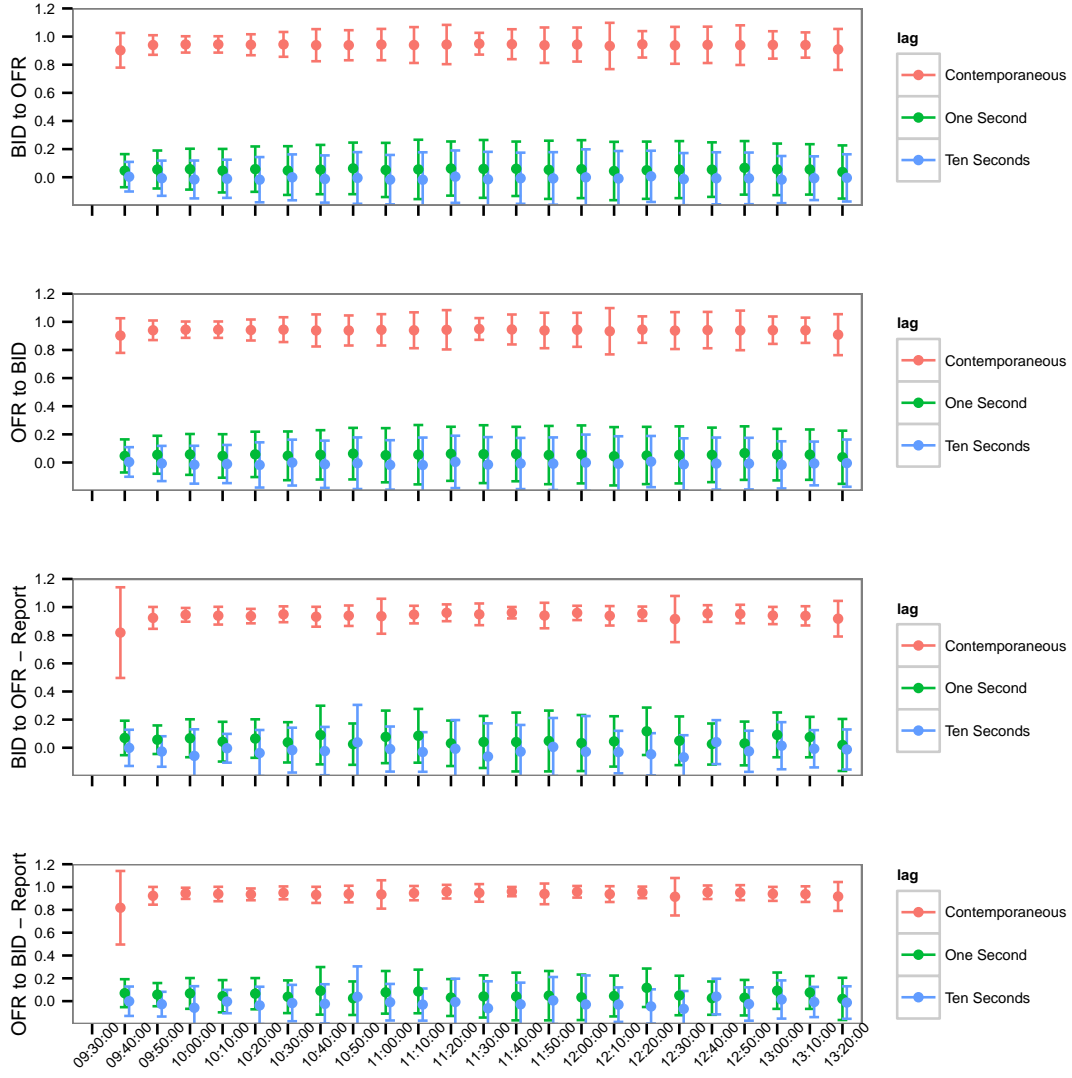


Figure 4: Time-Lagged Correlation between Revisions to Quotes (Bid-to-Ask & Ask-to-Bid) in Nearby and 1 Deferred Contracts

Mean correlations and one standard deviation error bars over all days in the top two plots. Mean correlations and one standard deviation error bars in the bottom two plots. Bid-to-Offer means revisions to the lagged nearby bid were correlated with the first deferred revisions to the offer, and Offer-to-Bid means revisions to the lagged nearby offer were correlated with the first deferred revisions to the bid.

References

- Aït-Sahalia, Y., P.A. Mykland, and L. Zhang. 2005. "How often to sample a continuous-time process in the presence of market microstructure noise." *Review of Financial Studies* 18(2):351–416.
- Akerlof, G.A. 1970. "The market for 'lemons': Quality uncertainty and the market mechanism." *The Quarterly Journal of Economics* 84(3):488–500.
- Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys. 2001. "The distribution of realized exchange rate volatility." *Journal of the American Statistical Association* 96(453):42–55.
- Avalos, F. 2014. "Do oil prices drive food prices? The tale of a structural break." *Journal of International Money and Finance* 42(0):253–271. Available at: <http://www.sciencedirect.com/science/article/pii/S0261560613001162>.
- Bagehot, W. 1971. "The only game in town." *Financial Analysts Journal* 27(2):12–14.
- Balcombe, K., A. Prakash, and others. 2011. "The nature and determinants of volatility in agricultural prices: An empirical study." *Safeguarding food security in volatile global markets*:89–110.
- Boehmer, E., J. Grammig, and E. Theissen. 2007. "Estimating the probability of informed trading-does trade misclassification matter?" *Journal of Financial Markets* 10(1):26–47. Available at: <http://www.scopus.com/inward/record.url?eid=2-s2.0-33846279422&partnerID=40&md5=864042ba646f1d94703dccc526ad2da8>.
- Cheng, I.-H., and W. Xiong. 2013. "The financialization of commodity markets." Working paper series No. 19642, National Bureau of Economic Research. Available at: <http://www.nber.org/papers/w19642>.
- Copeland, T.E., and D. Galai. 1983. "Information effects on the bid-ask spread." *The Journal of Finance* 38(5):1457–1469.
- Easley, D., and M. O'Hara. 1987. "Price, trade size, and information in securities markets." *Journal of Financial economics* 19(1):69–90.
- Easley, D., and M. O'Hara. 1992. "Time and the process of security price adjustment." *The Journal of Finance* 47(2):577–605.
- Easley, D., S. Hvidkjaer, and M. O'Hara. 2002. "Is information risk a determinant of asset returns?" *The Journal of Finance* 57(5):2185–2221.
- Easley, D., N.M. Kiefer, and M. O'Hara. 1997. "One day in the life of a very common stock." *Review of Financial Studies* 10(3):805–835.
- Easley, D., M.M.L. de Prado, and M. O'Hara. 2012. "The volume clock: Insights into the high-frequency paradigm (digest summary)." *Journal of Portfolio Management* 39(1):19–29.
- Gardebroek, C., and M.A. Hernandez. 2013. "Do energy prices stimulate food price volatility? Examining volatility transmission between {uS} oil, ethanol and corn markets." *Energy Economics* 40(0):119–129. Available at: <http://www.sciencedirect.com/science/article/pii/S0140988313001266>.
- Gilbert, C.L., and C.W. Morgan. 2010. "Food price volatility." *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 365(1554):3023–3034.
- Glosten, L.R., and P.R. Milgrom. 1985. "Bid, ask and transaction prices in a specialist market with heterogeneously informed traders." *Journal of Financial Economics* 14(1):71–100.
- Hasbrouck, J. 2006. *Empirical market microstructure: The institutions, economics, and econometrics of securities trading*. Oxford University Press.

- Hasbrouck, J. 2015. "High frequency quoting: Short-term volatility in bids and offers." New York University. Available at: <http://ssrn.com/abstract=2237499>.
- Hasbrouck, J. 2013. "High frequency quoting: Short-term volatility in bids and offers." *Available at SSRN 2237499*.
- Henderson, B.J., N.D. Pearson, and L. Wang. 2015. "New evidence on the financialization of commodity markets." *Review of Financial Studies* 28(5):1285–1311. Available at: <http://rfs.oxfordjournals.org/content/28/5/1285.abstract>.
- Irwin, S.H., and D.R. Sanders. 2011. "Index funds, financialization, and commodity futures markets." *Applied Economic Perspectives and Policy* 33(1):1–31. Available at: <http://aepp.oxfordjournals.org/content/33/1/1.abstract>.
- Irwin, S.H., and D.R. Sanders. 2012. "Testing the masters hypothesis in commodity futures markets." *Energy Economics* 34(1):256–269. Available at: <http://www.sciencedirect.com/science/article/pii/S0140988311002362>.
- Kumar, P., and D.J. Seppi. 1994. "Information and index arbitrage." *Journal of Business* 67(4):481–509.
- Lee, S.S., and P.A. Mykland. 2008. "Jumps in financial markets: A new nonparametric test and jump dynamics." *Review of Financial Studies* 21(6):2535–2563.
- 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 and Policy* 36(3):504–526.
- Mallory, M.L., S.H. Irwin, and D.J. Hayes. 2012. "How market efficiency and the theory of storage link corn and ethanol markets." *Energy Economics* 34(6):2157–2166. Available at: <http://www.sciencedirect.com/science/article/pii/S0140988312000552>.
- O'Hara, M. 1995. *Market microstructure theory*. Blackwell Cambridge, MA.
- Peterson, P. 2015. "How will closing the trading pits affect market performance." *farmdoc daily* 5(40).
- Prakash, A., C.L. Gilbert, and others. 2011. "Rising vulnerability in the global food system: Beyond market fundamentals." *Safeguarding food security in volatile global markets*:45–66.
- Serra, T., and D. Zilberman. 2013. "Biofuel-related price transmission literature: A review." *Energy Economics* 37(0):141–151. Available at: <http://www.sciencedirect.com/science/article/pii/S0140988313000388>.
- Trujillo-Barrera, A., M. Mallory, P. Garcia, and others. 2012. "Volatility spillovers in uS crude oil, ethanol, and corn futures markets." *Journal of Agricultural and Resource Economics* 37(2):247.
- Vacha, L., K. Janda, L. Kristoufek, and D. Zilberman. 2013. "Time–frequency dynamics of biofuel–fuel–food system." *Energy Economics* 40(0):233–241. Available at: <http://www.sciencedirect.com/science/article/pii/S0140988313001424>.
- Vega, C. 2006. "Stock price reaction to public and private information." *Journal of Financial Economics* 82(1):103–133. Available at: <http://www.scopus.com/inward/record.url?eid=2-s2.0-33748041619&partnerID=40&md5=1eb71f6d3df8978c80f959991cfc4348>.
- Wang, X. 2014. *Price volatility and liquidity cost in grain futures markets*. PhD thesis. University of Illinois.
- Yan, Y., and S. Zhang. 2012. "An improved estimation method and empirical properties of the probability of informed trading." *Journal of Banking & Finance* 36(2):454–467. Available at: <http://www.sciencedirect.com/science/article/pii/S0378426611002433>.