

THREE ESSAYS ON COMMODITY MARKETS' QUALITY

BY

QUANBIAO SHANG

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Agricultural and Applied Economics
in the Graduate College of the
University of Illinois Urbana-Champaign, 2022

Urbana, Illinois

Doctoral Committee:

Professor Maria Teresa Serra Devesa, Chair
Professor Emeritus Philip Garcia
Professor Scott H. Irwin
Associate Professor Mindy Mallory, Purdue University

ABSTRACT

Agricultural commodity prices adjust to ever-changing factors including weather, technology, and policy. Prices signal which goods are abundant and which are scarce and generate responses in production, consumption, and trade. Frequent price changes increase business risks and drive the demand for hedging. In this context, derivatives markets play a key role in price discovery, price forecasting, and risk management. Since the introduction of electronic trading, U.S. agricultural commodity futures markets have undergone relevant changes that have raised market quality concerns. A reduced market quality may have implications for hedging effectiveness as well as for efficiency in price discovery and forecasting. This dissertation centers around commodity markets' quality and their roles in price discovery and risk management.

The first essay focuses on hidden liquidity in agricultural futures markets. Like many exchanges, the Chicago Mercantile Exchange (CME) allows traders to conceal part of their limit orders, known as a hidden-limit order (HLO) or iceberg order. This type of order is well known for attracting traders who cannot compete in the market in terms of speed. However, with HLOs, market participants have incomplete knowledge of the limit order book. We assess the effect of this lack of transparency on market quality during a period of high volatility. We focus on corn and live cattle futures markets. Our conservative estimates indicate that HLOs represent more than 10% (20%) of the total volume in corn (live cattle) futures market. The findings show that the existence of HLOs improves market quality in multiple dimensions: driving trading volume while reducing market volatility and enhancing market liquidity. Our results are critical for regulators and exchanges as they are supportive of a certain degree of opacity. They are also indicative that HLOs can protect traders who have speed disadvantages in the era of fast trading.

The second essay investigates the existence of persistent profit opportunities in futures markets, which raises doubts on the quality of futures markets and challenges the market efficiency hypothesis. Specifically, with more than 30 years of data, we investigate Spread Time Series Momentum (STSM) in 22 U.S. commodity futures markets. First, we assess whether past spread returns can predict future returns, a necessary condition for the existence of momentum. We find predictability to be very weak after correcting for the issues affecting prior research. Second, we implement STSM-based investment strategies. We compare STSM profits for individual markets and portfolios to profits generated by a simple long-only benchmark strategy that does not require any predictability. STSM does not generate returns statistically different from the benchmark trading strategy, with both strategies generating very low or negative returns. For the momentum to outperform the benchmark strategy, predictability should be three times larger than observed from real data, but would entail substantial downside risk. In sum, the empirical evidence indicates that returns from STSM-type strategies are illusive for the commodities and period studied. The inclusion of unrealizable roll yield generates the illusion of profitable STSM trading strategies in previous research.

The third essay investigates the role of U.S. agricultural commodity swap markets in price discovery and risk management. While swap contracts have become increasingly relevant financial derivatives in the past two decades, little is known about these markets. Using real-time released swap trade data for corn and soybeans from 2018 to 2019, we investigate swap markets from three different perspectives. First, we provide a descriptive analysis of the characteristics of swap markets relative to their futures market counterpart. We find that swap volume accounts for more than 16% (19%) of the total volume in corn (soybean) nearby futures market. Second, we investigate price discovery between futures and swap markets. Findings from our intraday analysis

suggest that swaps and futures markets' prices are bonded by an equilibrium relationship and their respective information shares are commensurate with their respective volume shares. Efficient price discovery in financial derivatives is at the core of hedging effectiveness, which we investigate in the third part of the article. Results suggest that futures contracts are superior to swaps in short hedging horizons. However, the two instruments are equally effective in hedging cash price risks for hedging horizons 35 (50) days or longer in corn (soybean) market.

ACKNOWLEDGEMENTS

First and foremost, I would express my sincere gratitude to my advisor Professor Teresa Serra for her valuable advices and patience during my Ph.D. career. Her immense knowledge and plentiful experience have encouraged me in my academic research and daily life. Secondly, many sincere thanks to Professor Philip Garcia, who spends enormous amount of time in helping me since I became a graduate student. He is a “scholars’ scholar” and I desire to learn much more from him in the coming years. In addition, I wish to present my gratitude to Professor Scott Irwin and Professor Mindy Mallory. Professor Irwin always has his door open to me for questions and support, and Professor Mindy Mallory guided me into the field of futures market and consistently giving me valuable advices. Lastly, my appreciation goes out to my family and friends for their love and encouragement through my studies.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION	1
CHAPTER 2: LOOKING UNDER THE SURFACE: AN ANALYSIS OF ICEBERG ORDERS IN THE U.S. AGRICULTURAL FUTURES MARKETS	5
CHAPTER 3: RIDE THE TREND: IS SPREAD MOMENTUM A RELIABLE TRADING INDICATOR IN THE U.S. COMMODITY MARKETS?	52
CHAPTER 4: THE ROLE OF SWAP CONTRACTS IN THE U.S. AGRICULTURAL COMMODITY MARKETS	97
CHAPTER 5: CONCLUSIONS	148

CHAPTER 1

INTRODUCTION

Agricultural commodity prices adjust to ever-changing factors including weather, technology, and policy. Prices signal which goods are abundant and which are scarce, causing responses in production, consumption, and trade. Frequent price changes increase business risks and drive the demand for hedging. In this context, derivatives markets play a key role in price discovery, price forecasting, and risk management. Since the introduction of electronic trading, U.S. agricultural commodity futures markets have undergone relevant changes that have raised market quality concerns. A reduced market quality may have implications for hedging effectiveness as well as for efficiency in price discovery and forecasting. This dissertation centers around commodity markets' quality and their roles in price discovery and risk management in the era of electronic trading. This dissertation is timely and identifies three circumstances in which market quality can be compromised that have not been previously investigated. These include futures' market design allowing for a certain degree of opacity, existence of persistent arbitrage opportunities in the futures markets, and the role of the rather opaque swaps market in forward trading and its interactions with futures markets.

The first essay examines the presence of hidden limit orders (HLOs) in agricultural futures markets and their consequences for market quality. Like many exchanges the Chicago Mercantile Exchange allows traders to conceal part of their limit orders, known as a HLO or iceberg order. The lack of transparency surrounding hidden liquidity (HL) has generated concerns about its magnitude and consequences for market quality (Boulatov and George 2013). The market quality implications of HLOs are important for the market and for the traders who hide. We use market

limit order book data along with transactions in the corn and live cattle futures markets. To identify the presence of HLOs, we implement an algorithm that centers on the difference between transactions size and the change in limit order book depth. Although HLOs are not directly observable, under the current order matching mechanism, this technique allows us to capture the presence of HLOs. To identify the HLO effects on market quality, we create a control-and-treatment type method to examine how market quality differs with and without the presence of HLOs. We measure market quality within the 1-minute *ex-post* interval following an HLO execution through regression analysis. For the first time, we quantify the impacts of HLOs on price volatility, spreads, and volume.

The second essay investigates the futures market calendar spread time series momentum (STSM) in U.S. commodity futures markets. Some previous researchers have argued that trading strategies based on STSM can deliver significant returns (Szymanowska et al. 2014; Boons and Prado 2019), which, if true, is at odds with the efficient market hypothesis. These arguments however, are based on analyses that do not exclude the unrealizable futures contract roll yield and are also affected by other empirical and statistical issues that may lead to misleading results. In this article, we examine whether STSM leads to high profitability in both spread (STSM-spread) and outright (STSM-outright) investment strategies. We use a sample of 22 U.S. commodity futures markets observed from 1986 to 2019. For the first time in the momentum literature, we adopt a comprehensive approach that corrects for both the empirical design and econometric issues affecting previous research. First, we assess whether past spread returns can predict future returns, a necessary condition for the existence of momentum. Second, we implement STSM-based investment strategies. We compare STSM profits for individual markets and portfolios to profits generated by a simple long-only benchmark strategy that does not require any predictability.

The third essay explores the once opaque agricultural swap markets. Over the last couple of decades, swap contracts have become increasingly relevant in commodity markets (Irwin and Sanders 2011; Mixon, Onur, and Riggs 2018). Research on swap contracts has however been very limited due to the opacity that characterized this market for a long time (Mixon, Onur, and Riggs, 2018). Despite increased transparency and in contrast to attention received from regulatory agencies, research on agricultural commodity swap markets is scarce. With the real-time released corn and soybean swap trade data from 2018 to 2019, we investigate swap markets from three different perspectives. First, we provide for the first time a general descriptive analysis on how these markets compare in size, trade volume, and patterns with their futures counterparts. We then shed light, also for the first time, on their role in market price discovery. To make swap and futures prices comparable, we use the cost-of-carry model for storable commodities to discount swap prices into a nearby futures-equivalent price. Lastly, we investigate swaps effectiveness in hedging cash positions relative to futures markets. We choose to study price discovery and hedging effectiveness of swap contracts as they are two fundamental roles of financial derivatives markets (Working, 1962; Garbade and Silber, 1983).

1.1.References

- Boons, M. and Prado, M.P., 2019. Basis-momentum. *The Journal of Finance*, 74(1), pp.239-279.
- Boulatov, A., & George, T. J. (2013). Hidden and displayed liquidity in securities markets with informed liquidity providers. *The Review of Financial Studies*, 26(8), 2096-2137.
- Garbade, K.D. and Silber, W.L., 1983. Price movements and price discovery in futures and cash markets. *The Review of Economics and Statistics*, pp.289-297.
- Irwin, S.H. and Sanders, D.R., 2011. Index funds, financialization, and commodity futures

- markets. *Applied Economic Perspectives and Policy*, 33(1), pp.1-31.
- Main, S., Irwin, S.H., Sanders, D.R. and Smith, A., 2018. Financialization and the returns to commodity investments. *Journal of Commodity Markets*, 10, pp.22-28.
- Mixon, S., Onur, E. and Riggs, L., 2018. Integrating swaps and futures: A new direction for commodity research. *Journal of Commodity Markets*, 10, pp.3-21.
- Szymanowska, M., De Roon, F., Nijman, T. and Van Den Goorbergh, R., 2014. An anatomy of commodity futures risk premia. *The Journal of Finance*, 69(1), pp.453-482.
- Working, H., 1962. New concepts concerning futures markets and prices. *The American Economic Review*, 51(2), pp.160-163.

CHAPTER 2

LOOKING UNDER THE SURFACE: AN ANALYSIS OF ICEBERG ORDERS IN THE US AGRICULTURAL FUTURES MARKETS

2.1. Introduction

Current U.S. agricultural commodity futures contracts are traded on order-driven electronic markets based on open limit order books (LOBs). Trading activities are observed in real-time by market participants, which contributes to market transparency. However, not all quoting activity in the market can be fully observed as exchanges allow traders to place hidden limit orders (HLOs) or iceberg orders. HLOs permit traders to specify the price and number of contracts they wish to buy or sell and how much of their intentions are displayed in the LOB. The portion of the limit order they choose to display is the peak size. The remaining portion is hidden to the public and automatically replenishes the peak size each time it is executed, until the entire HLO is executed or cancelled. The newly replenished quantity, however loses execution priority to previously displayed orders at the same price. This increases the time it takes to fill these orders and reduces the likelihood that they can be fully executed at the specified price.

The lack of transparency surrounding hidden liquidity (HL) has generated concerns about its magnitude and consequences for market quality (Boulatov and George 2013). Current knowledge about HL comes from analyses of stock market data (Degryse 1999; Hasbrouck and Saar 2002; Pascual and Pardo 2012; Frey and Sandas 2017; Jain and Jain 2017), which place HL between 10 and 50 percent of total market liquidity and conclude that HLOs have, on balance, a positive effect on markets. However, these results may not apply to agricultural commodity futures markets, which have a different purpose and involve different participants. Our study

identifies the patterns and impacts of HLOs on intraday market quality in the U.S. agricultural futures markets.

Understanding the relevance of our analysis requires discussion of who is hiding and why. Traders may be wary of exposing large limit orders in the LOB since this may trigger price movements against their interests (Esser and Monch 2007; Bloomfield et al. 2015; Anand, Chakravarty, and Martell 2005; Moinas 2010). For example, when new valuable information becomes available, liquidity providers risk being adversely selected by fast informed traders if they delay updating their quotes. To reduce the risk of being outpaced by fast informed traders, slow liquidity providers can either rely on advanced technologies that allow fast quote updates or alternatively, they can hide their positions to overcome their speed disadvantage. In contrast, fast liquidity providers are less likely to use HLOs since these positions have a lower probability of being executed than regular limit orders at the same price level, which may reduce profits (Gao 2015). Roşu (2019) explains how fast, informed traders usually make a profit at the expense of slower informed traders by quickly betting on their information and moving the market price. After doing so, they quickly reverse their positions, usually matched against traders who respond more slowly to the initial information. As a result, slower traders get their positions filled at a much less attractive price. Hence, in the era of fast trading, hiding positions may help both slow informed and uninformed traders to overcome speed disadvantages. Agricultural commodity futures markets involve hedgers, commercials and other market participants who tend to be more information-driven and far less competitive in speed than high-frequency traders. This collective has expressed dissatisfaction with a current market design that concedes speed advantage to fast traders (USDA 2012). As a result, the option to hide may be crucial to protect these slower traders and the market's overall health (Working 1970; Boulatov and George 2013).

Despite their name, HLOs are detectable through mismatches in the LOB every time the peak size is executed and replenished. This may trigger responses from traders and affect market conditions. Slow traders will place HLOs if they can be filled without triggering increased market volatility and wider spreads after each partial execution. This would increase their chances of not being filled or being filled at an unfavorable price if fast traders take the other side in anticipation of an impending price movement. Holders of HLOs are further interested in driving volume to the market so that the remaining unexecuted part of the order can be filled as well. Hence, the evolution of market quality (volatility, spread and volume) after the peak size is executed is key to assess whether traders hide successfully and thus whether HLOs can help address their speed disadvantage.

This article investigates the patterns and impacts of hidden liquidity (HL) in agricultural commodities markets for the first time. The analysis consists of two different parts. We first use a HLO detection algorithm developed and employed in previous studies (Pardo and Pascual 2012; Frey and Sandas 2017) that offers a conservative measure of HL by identifying only the portion of the HLOs that is matched against an incoming order. In the second part of the analysis, we use regression techniques to identify the consequences of HL on market quality. Previous approaches have investigated the impacts of HLOs on market quality by using either theoretical or experimental approaches (Boulatov and George 2013; Kovaleva and Iori 2015; Bloomfield et al. 2015), or empirical approaches that focus on the sign, but not the magnitude of these impacts (Pascual and Pardo 2012). Others are focused on identifying the reasons that lead traders to hide (Frey and Sandas 2017) and assume that HLOs will have positive or negative market impacts, depending on whether uninformed or informed traders place them. This is the first study that quantifies the impacts of HLOs using market data and a regression analysis that measures the

impacts of HLO on price volatility, spreads and volume.

To provide insights into the use and consequences of HLOs, we consider two markets with substantially different degrees of liquidity: the actively traded corn futures market and the lightly traded live cattle futures market (Irwin and Sanders 2012). In addition, we study HLOs in contracts with different maturities (nearby and first deferred), which are also characterized by different levels of activity and volatility. We use Market Depth (MD) data from the CME Globex trading platform from November 2015 to May 2017, which corresponds to a particularly volatile time in both markets when both nearby corn and live cattle futures prices declined by roughly 30%.

We find that HL represents more than 10% (20%) of the total volume in the corn (live cattle) futures market. Findings suggest that the existence of HLOs improves market quality in multiple dimensions: increasing trading volume (7% in corn and 5% in live cattle) while reducing market volatility by at least 5% and reducing the relative trading spread by least 11%. These results suggest that hidden liquidity can successfully serve the purpose of slow traders and improve market quality. Also, hidden liquidity can provide a signal for slower traders of subsequent market quality conditions, allowing them to establish market positions. Our results are relevant for market design and surveillance. For the exchange, it is important to assess the impacts of HLOs in order to make more informed decisions for future market design. For slow traders and other market participants, understanding HL's magnitude, patterns, and impacts is of relevance to manage order submission strategies.

2.2. Data

We use day trading hours from the CME Group Market Depth (MD) data from November 23, 2015 to May 31, 2017 for the corn and live cattle futures markets. A description of the dataset is provided in the supplementary results. The analysis is based on both the nearby and first deferred

contracts, where the nearby (deferred) is the first (second) closest to expiration contract. The first deferred contract replaces the nearby contract on the first day of the nearby maturity month.¹

Table 2.1 provides summary statistics for the contracts studied. Both contract volume transacted and the number of transactions per day are substantially higher in the nearby liquid contracts (first and second columns). Given the much larger trading volume in the corn market relative to the live cattle market, choosing these markets enriches the study by providing evidence on the cross-market differences of HLOs. The daily quoted bid-ask spread (BAS) is shown in the fourth column. In the corn market, the BAS for the nearby and deferred contracts are similar and only slightly larger than the minimum tick size (0.25 cents/bu.). The BAS in the live cattle market is 0.040 cents/lb. in the nearby, which is about 50% higher than the minimum tick size (0.025 cents/lb.) and nearly two ticks in the first deferred contract. To describe liquidity conditions, we also report the top-level quoted depth in table 2.1, calculated as the average of the top-level ask and bid depth. In the corn nearby market, the depth is more than 200 contracts on either side of the market, which contrasts with the 7 contracts in the live cattle market. The depth declines significantly in the first deferred contract, with 97 contracts being offered at the best price in the corn market and only 4 contracts in the live cattle market. For both markets, we measure the daily volatility by the standard deviation of quote midpoints. Volatility in the corn (live cattle) nearby contract is 4.6% (6.1%) higher than the deferred contract.

2.3. Hidden Liquidity Patterns

Hidden limit orders are special limit orders that have all features of a regular limit order plus the

¹ Due to the uncertainty of the timing of new crop delivery, the September corn futures contract is lightly traded (Smith 2005; Wang et al. 2014) and is not included in the analysis.

peak size, which is the specified size displayed to the market. The non-displayed portion of the limit order is hidden awaiting execution. After the peak is filled, a new quantity equal to the specified peak is automatically replenished to the displayed depth. The replenished order loses its original time precedence and the time stamp when it is replenished defines its new execution priority. This process repeats until the entire HLO is exhausted or canceled. If an aggressive incoming order only depletes part of the displayed peak size, no auto-replenishment is triggered.

2.3.1. Hidden Liquidity Detection Algorithm

While CME makes the displayed LOB available to the public, HLO submissions and cancellations are not identified in the CME MD data. However, researchers have developed a reliable algorithm to identify HL that requires the displayed HLO volume to be matched against incoming market orders (Pardo and Pascual 2012; Hautsch and Huang 2012; Frey and Sandas 2017). Under the current matching mechanism, any difference between market-order size and the change in quoted depth defines the HLO auto-replenishment. Figure 2.1 provides an example of how HLOs are detected. In the figure, the quoted size after a trade takes place (post-trade quoted size) does not match the difference between the pre-trade quoted size and the trade size due to auto-replenishment of the peak size. With a hypothetical LOB, table 2.7 in the supplementary results further demonstrates how the algorithm works by presenting three scenarios in which HLO can be identified. The algorithm identifies the executed portion of an HLO but not the remaining unexecuted part, thus offering conservative estimates. Frey and Sandas (2017) show the high accuracy of this algorithm with a dataset that contained actual HLO submissions.

2.3.2. Detection Algorithm Results

Results of the HLO detection algorithm are shown in table 2.2.² The daily trading volume matched against HL is 11.3% in the nearby contract and 17.8% in the first deferred in the corn market. In the live cattle market, the HLO ratio is twice as large relative to the corn market, with 23.1% in the nearby and 26.8% in the first deferred contract. Results are consistent with previous research findings that HLOs are more prevalent in markets characterized by low trading activity and wider bid-ask spreads (Bessembinder et al., 2009; De Winne and D'Hondt, 2007). Table 2.3 also provides the HLO ratio based on the proportion of transactions. About 11% (21.5%) of transactions are matched at least partially against HL in the corn (live cattle) nearby market. This number in the deferred contract is 15.3% (26.7%) in the corn (live cattle) market. The HLO trade ratio is slightly smaller than the HLO volume ratio because the average trade size-matched against HL is much larger than trades involving no HL.

The last two columns in table 2.3 provide insights into the effect of trade size on the appearance of HL. In the nearby corn market, the average transaction size that triggers an HLO replenishment (19 contracts/trade) is considerably larger than the average non-HLO transaction size (5 contracts/trade). A similar pattern emerges in the live cattle nearby market. However, the size difference between transaction types is much smaller than in the corn market (4 vs 2 contracts/trade). HLOs are revealed more easily and frequently in the live cattle market with scarcer liquidity. The same patterns characterize the deferred contracts, which have overall smaller trade sizes. As discussed earlier, an only complete exhaustion of the display portion

² Based on the three different scenarios that HLO can be revealed, we report the daily average volume estimates separately for each scenario in table 2.8 (supplementary results).

results in auto-replenishment. Hence, it is not surprising to see that larger orders tend to trigger HLO auto-replenishment more often in both nearby and deferred contracts.

Overall, our results are within the range of estimations of HL by previous research which concludes that HL represents from 10 to 50 percent of the total market liquidity in different financial markets. Degryse (1999), Hasbrouck and Saar (2002), Pascual and Pardo (2012) and Frey and Sandas (2017) place their numbers between 10 and 15 percent. Tuttle (2006) reports that HL accounts for 22% of the total market depth in the NASDAQ market. However, Bessembinder, Panayides, and Venkataraman (2009) use a Euronext-Paris stocks sample and find that HLOs can represent more than 50% of total volume in the less liquid markets. With historical data that contains accurate HLO information, DeWinne and D'Hondt (2007) derive analogous results in the Euronext market. On average, HL accounts for 45% of the total depth in the top 5 levels of the LOB.

Figure 2.2 presents the daily HLOs as a percent of total volume for the entire period. On most days, the corn nearby contract HLO ratio fluctuates around 10%, while the HLO ratio for the first deferred contract is consistently higher.³ Near contract expiration, the HLO ratio in the deferred contract falls below the nearby contract, whose ratio spikes. Wang et al. (2014) show that the BAS sharply increases as expiration approaches and volume vanishes. As a result, liquidity provision becomes riskier and motivates increased HLOs (Bessembinder et al. 2009). With the switch to the first deferred contract, the HLO ratio spike disappears. Aside from the contract-rolling periods, the HLO ratio is relatively stable. However, in summer 2016, a

³ The HLO ratio does not change appreciably across different trading hours and thus shows a flat pattern throughout the day.

systematic increase in HLO emerged in both corn contract series. In this period, the corn price suffered a sharp decline from \$4.3/bu. to \$3.2/bu. And overall volatility increased. In periods with higher volatility and uncertainty, the cost of posting limit orders increases, and liquidity providers demand higher compensation (Copeland and Galai 1983). Not surprisingly, HLOs increase simultaneously. The daily HLO ratio shows a similar pattern in the live cattle market. In addition to the higher HLO ratios, we see more extended periods when the first deferred contracts have a lower HLO ratio than the nearby contract. This is consistent with Hu et al. (2020), who show that the nearby live cattle futures contract loses its volume share dominance to deferred contracts as early as 5-6 weeks before expiration, and with Bessembinder et al. (2009), who show increased HLOs being a response to low trading activity.

Table 2.3 provides an overview of how daily HLO ratios vary on the USDA announcement days compared to the rest of the trading days. For corn, we consider the three most important USDA grain reports: The World Agricultural Supply and Demand Estimate Report (WASDE), the Crop Production report (CP), and the Grain Stocks (GS) report. Following Isengildina et al. (2006), we focus on the Cattle on Feed report as the most critical USDA report in the live cattle market.

We report the mean of the daily HLO volume ratio and use Welch's t-test to compare the mean values between the two types of trading days. In all markets and contracts, we see that HL tends to be more prevalent on USDA announcement days than on non-announcement days. The HLO volume ratio in the nearby contract is about 20% higher on USDA announcement days in both markets. Similar patterns are identified in the deferred contracts. With higher volatility and added adverse selection risk on announcement days, it is not surprising to see higher HLO ratios as liquidity providers and slow informed traders protect their trading activity.

2.4. HLO Impacts on Market Quality

Given the incidence of HL in the corn and live cattle futures markets, we anticipate the consequences of HL on observable market quality characteristics will emerge. Our empirical strategy consists of a regression analysis that measures *ex-post* market conditions after transactions involving HLO relative to non-HLO trades. By including a dummy variable that identifies HLO, we measure whether there are discernable differences. To permit clearer identification of the HLO effects, we resample the data. We only include transactions that have neither prior nor subsequent HLO-trades in the 1-minute *ex-ante* and *ex-post* time intervals. We pair each HLO-trade with another non-HLO-trade within the same trading day, at least 2-minutes apart from the HLO trade. This ensures that the 1-minute *ex-ante* and *ex-post* market conditions of the paired trades do not overlap. The pairing algorithm is similar to the procedure used by Pardo and Pascual (2012). On each day, the 1-minute *ex-ante* BAS and volatility are separated into six percentiles for all transactions. For an HLO and non-HLO trade to be paired, both the pre-trade volatility and BAS of the two trades need to be in the same percentile. Imposing such a matching algorithm creates an experiment-like study where HLO trades and non-HLO trades are only compared if they have a similar preceding market environment. When multiple non-HLO-trades can be matched with an HLO-trade, we select the trade closest in time. Restricting the HLO and non-HLO trade pair to occur within the same day ensures the two trades occur under similar macroeconomic and fundamental conditions.

The market quality implications of HLOs are important for the market and for the traders who hide. We measure market quality within the 1-minute *ex-post* interval following an HLO execution. The 1-minute time interval choice is made based on several considerations. First, the number of observations decreases as we lengthen the time interval. Because the subsequent effect

of HLO is rather short-lived (Pardo and Pascual 2012), the impact of HLOs in a longer time interval becomes ambiguous and difficult to identify. In addition, in the less liquid live cattle, finer intervals can produce time intervals with no (or too few) trade occurrences.

The three dimensions of market quality identified earlier are measured *ex-post* in the minute following the transaction: volatility (Vlt_{t+1}), transaction volume (Vlm_{t+1}), and relative spread (Lqd_{t+1}). Volatility (Vlt_{t+1}) is the coefficient of variation of quote midpoints. The total trading volume (Vlm_{t+1}) is normalized by the market's average daily trading volume for the same month in which it occurs. The third variable reflects the spread, which is normalized by market depth following Chordia et al. (2001), $Lqd_{t+1} = (\%BAS/\$Depth)$, where $\%BAS$ is the quoted bid-ask spread divided by the quote midpoint expressed as a percent, and $\$Depth$ is the average of the best ask depth times ask price and best bid depth times bid price. This measure allows us to capture the capacity of the market better to absorb transactions. The smaller this index, the thicker and balanced is the book on both the buy and sell sides.

The following regressions are estimated for each market (corn and live cattle) and contract (nearby and first deferred):

$$Vlt_{t+1} = f(BAS_{t-1}, HLO_t, OAgr_t, BfrAnn, AftAnn_5m, AftAnn, Opn, Cls, Vlt_{t-1}, Vlm_{t-1}) \quad (2.1)$$

$$Vlm_{t+1} = f(BAS_{t-1}, HLO_t, OAgr_t, BfrAnn, AftAnn_5m, AftAnn, Opn, Cls, Vlt_{t-1}, Vlm_{t-1}) \quad (2.2)$$

$$Lqd_{t+1} = f(HLO_t, TSize_t, OIB_{t-1}, BfrAnn, AftAnn_5m, AftAnn, Opn, Cls, Vlt_{t-1}, Vlm_{t-1}, Lqd_{t-1}) , \quad (2.3)$$

where t denotes the transaction time. Volatility (Vlt_{t+1}), transaction volume (Vlm_{t+1}), and the relative spread (Lqd_{t+1}) are all expressed in logs to facilitate interpretation. To assess the market effects of HLO in all three equations, we include a dummy variable that distinguishes between HLO orders ($HLO_t = 1$) and non-HLO orders ($HLO_t = 0$). With the HLO detection strategy, we only observe the proportion of the HLO that is matched against incoming market orders. The larger a market order, the larger the revealed HLO size due to the nature of the limit order book matching algorithm (Frey and Sandas 2017). To avoid potential multicollinearity between the revealed HLO size and the trade size, we use a dummy variable to reflect the presence of HLOs. To measure the aggressiveness ($OAgr_t$) of a market order, we subtract the market order size from the prevailing matching-side depth, normalized by the average same-side market depth in the same month. We accommodate persistence in volatility, volume, and the spread (e.g., Andersen and Bollerslev 1997) by incorporating a 1-minute lag of the dependent variable in each equation (Vlt_{t-1} , Vlm_{t-1} , and Lqd_{t-1}). BAS_{t-1} (equations 2.1 and 2.2) is the average bid-ask spread normalized by the median market BAS during the month, which signals the market quality 1-minute before each transaction. The higher the BAS_{t-1} , the more likely we observe the reduced market quality after the transaction (Stoll 1978; Ho and Stoll 1983).

Notice that our relative spread (Lqd_{t+1}) equation 2.3 differs in the variables included. Since Lqd_{t+1} reflects both the BAS and the quoted depth and is used as an explanatory variable in lagged form. We do not use the BAS and order aggressiveness variables in (2.3) to prevent multicollinearity. Instead, we use transaction size ($TSize_t$) measured in the number of contracts, normalized by the market's average daily trading volume of the same month. $TradeSize_t$ provides relevant information to the market as large aggressive orders that exhaust a large amount (or all) of the top-level depth, increase adverse selection costs and are likely to have a

large impact on market quality (Easley and O'Hara 1987). Order imbalance (OIB_{t-1}) can be useful to infer the probability of informed trading and subsequent liquidity depletion (Easley et al. 2012). We measure order imbalance as the absolute value of the difference between total buys and sells in the number of contracts normalized by the market's average daily trading volume of the same month.

Several variables are also included to account for the public release of market information and conditions, which can vary systematically within the day. Earlier studies show relevant changes in market conditions both before and after USDA public announcement shocks (Lehecka et al., 2014; Adjemian and Irwin, 2018; Couleau et al., 2020). In the corn market, WASDE, GS, and PP announcements are released during normal trading hours. We capture their effects using three dummy variables to identify whether, on an announcement day, the transaction occurs before (*BfrAnn*), in the first five minutes (*AftAnn_5m*), or after the first five minutes (*AftAnn*) following the release of public information on the release day. *AftAnn* is based on previous research results identifying a lengthier price discovery when the reports are released in the middle of the trading session (Joseph and Garcia 2018). The Cattle on Feed report is released in the live cattle market after the market close during the sample period of study. The *BeforeAnn* dummy here includes the entire day trading session before report release. Also, in the live cattle equations, we only consider the *AfterAnn_5m* dummy. We did not consider a longer interval because previous research has shown that when the release occurs outside normal trading hours, price discovery occurs very quickly once the market reopens (Lehecka et al. 2014).

Existing studies have also identified that market conditions vary during trading hours. For example, trading activity and volatility are generally higher during market opening and closing than during mid-day hours (Shang et al., 2018; Joseph and Garcia, 2018), which is likely to affect

market quality. To capture these intraday patterns, we introduce dummy variables for the first (last) 30 minutes of market opening (closing), Opn (Cls).⁴ Random error terms are represented as $\varepsilon_{i,j}$, $\delta_{i,j}$, and $\mu_{i,j}$.

2.5. Empirical Results

In table 2.4, we present summary statistics of the selected samples. For the entire period of study, there are 4.2 (2.4) million transactions in the corn (live cattle) nearby market and 2.0 (1.5) million in the first deferred contract. After resampling, 69,526 (43,176) observations in the corn (live cattle) nearby contract, and 40,522 (22,580) observations in the corn (live cattle) deferred contract remain. The table also provides the average BAS, the size of the transactions, the average top-level quoted depth, and the number of subsample observations that coincide with USDA announcement days. Compared to the full sample BAS (table 2.1), only minimal differences emerge, indicating the selected samples represent the full sample.

Following Hayashi (2000), we use a GMM estimator to fit each regression model, yielding consistent and efficient estimates. In the presence of non-*iid* errors, the Newey and West (1994) method is applied to correct heteroscedasticity and autocorrelation in the standard errors. Tables 5 and 6 report coefficients and corresponding t-statistics for the corn and live cattle markets in the nearby and deferred series. In the supplementary results, we examine the robustness of the main results to samples with different *ex-ante* and *ex-post* intervals and by increasing the non-HLOs in the analysis. On balance, the results are robust to different empirical analysis designs.

HLO (HLO_t) has statistically significant impacts on all market quality measures. Results show HLO reduces the *ex-post* 1-minute market volatility by 4.4% (9.2%) in the corn nearby

⁴ $AftAnn_5m$ and Opn dummies do not overlap in the live cattle markets.

(first deferred) contract and by 6% (16.8%) in the live cattle nearby (first deferred) contract. These empirical findings contrast with Pascual and Pardo, who find an increase in ex-post market volatility.⁵ Still, they are consistent and lower than Kovaleva and Iori (2015) estimates based on a simulation study, who conclude that price volatility decreases as much as 23.8% when HL is allowed compared to markets with no HL. The buffer provided by HLOs helps to explain but may not constitute the only source of diminished volatility. As discussed later, regression results suggest that the emergence of HLOs triggers a response from latent liquidity demanders who become aggressive in trying to fill their positions at better immediacy costs due to the buffer offered by HLOs. However, their possible effect on volatility and spreads is counterbalanced by increased competition for liquidity provision, which reduces overall volatility. Pardo and Pascual (2012) and Frey and Sandas (2017) suggest that improved liquidity conditions are derived from liquidity providers trying to front-run HLOs once they find their existence by submitting displayed limit orders or more aggressive quote prices.

The size of the HLO_t coefficients differ visibly across markets and contracts in the Vlt_{t+1} equation. The absolute values of the coefficients are larger in the live cattle market than in the corn market for both the nearby and deferred. As discussed, the quoted depth in the corn market is much larger and the BAS is much tighter than in the cattle market. This results from a relatively higher level of competition in liquidity provision in corn relative to the live cattle market. As HLOs lose execution priority to displayed liquidity, the buffering effect of HLOs while still significant, becomes less substantial in a more liquid market. Consistently, we observe a stronger

⁵ Pascual and Pardo (2012) do not quantify the change in the ex-post market conditions and they only focus on the sign of the change.

HLO impact in the deferred contract in both markets than in the nearby. Recall that the average transaction size in the live cattle market is much closer to the top-level depth than corn (table 2.1). Hence, transactions in the live cattle market tend to exhaust multiple levels of depth more frequently, resulting in higher market quality impacts. As a result, the HL auto-replenishment process is essential to ensure adequate liquidity in the live cattle market. Thus, the role of the HLOs becomes more relevant in moderating volatility. Overall, HLOs decrease market volatility and this impact is stronger in less liquid markets.

Results in the volume (Vlm_{t+1}) equations provide evidence that the exposure of HLOs attracts latent liquidity demand. The HLO_t increases trading volume in the *ex-post* 1-minute interval by 7% (7.7%) in the corn nearby (deferred) contract. We find smaller estimates in the cattle market, 4.4% (5.5%). We can express the magnitude of volume increase in the number of contracts by multiplying the percentage change by the average 1-minute trading volume in each market.⁶ Results indicate that after exposure of HLOs, on average, there are 53 (6) more contracts traded in the corn (live cattle) nearby market and 31 (5) more contracts traded in the deferred markets. The finding is consistent with previous studies (e.g., Moinas 2010; Pardo and Pascual 2012; Frey and Sandas 2017), who conclude that liquidity demanders become more aggressive in submitting market orders to capture the opportunity of low transaction costs due to the buffer offered by HLOs. Such “searching and catching” HLO strategies may come from traders who closely monitor market order flows and respond quickly through automated trading programs. The proportion of automated trading is less in the live cattle market than in corn and the speed

⁶ The 1-minute average trading volume is 760 (148) contracts in corn (live cattle) market nearby series, and 402 (101) in the deferred series.

of order execution is also lower (Haynes and Roberts 2015, 2017, 2019). Thus, despite HLOs dampening volatility, its impact on promoting volume is more limited in the live cattle than in the corn market.

The third market quality indicator the relative spread (Lqd_{t+1}) captures the changes in both the BAS and the top-level quoted depth. Consistent with DeWinne and D'Hondt (2007) and Pardo and Pascual (2012), HLO coefficients are negative, indicating that liquidity conditions improve. As with reduced volatility, this may result from both a direct HLO buffer effect and traders being encouraged to supply additional liquidity with the emergence of HLOs. Results from equations 2.1 and 2.2 suggest that after HLOs are revealed, there is an increase in trading volume while volatility declines. This requires liquidity to improve, which we confirm with equation 2.3 results. On average, we see the relative spread improvement is more substantial in the live cattle market than in corn. We observe a 9.2% (1.2%) improvement in the corn nearby (deferred) contracts and a 15.4% (23.3%) improvement in the live cattle contracts. Our estimates are comparable to Kovaleva and Iori (2015), who find HLOs result in 15.2% liquidity improvement in their simulation study.

Overall, the HLO results are consistent with previous empirical studies that have found market participants do not seem to attribute relevant information to HLOs and the adverse selection issues are offset by the positive externalities of liquidity provision, leading to a rather neutral or slightly positive effect on market quality (Pardo and Pascual 2012; Frey and Sandas 2017). Theoretical or experimental approaches also minimize the possible negative impacts of informed HLOs. In Moinas (2010), the possibility of hiding motivates informed traders to submit limit orders, which results in increased informativeness of the limit order book, increased market efficiency, and increased risk of adverse selection. Bloomfield et al. (2015) find limited market

impacts of informed hidden orders on the market. Boulatov and George (2013) show that the possibility to hide increases the informational rents from informed traders and motivates these traders to provide rather than demand liquidity. This results in a narrower bid-ask spread, more informational mid-quotes and higher liquidity provision.

In general, the coefficient signs of the other variables in the analysis are consistent with the literature. Coefficient estimates suggest a positive relationship between BAS and the subsequent market volatility and a negative correlation between BAS and ensuing volume. These results are in line with previous studies that conclude that market volatility and trading volume are sensitive to liquidity costs (e.g., Wang et al., 2014; Wang and Yau, 2000). $OAg r_t$ parameters are negative in all volatility equations, suggesting that larger market orders tend to cause a higher level of volatility. However, order aggressiveness is not statistically significant in the live cattle market, which is likely related to the relatively small presence of aggressive orders. Within each market, the absolute value of the coefficient in the deferred contract is always larger than in the nearby contract, suggesting that nearby contracts are more resilient to aggressive orders than deferred contracts. As expected, the more aggressive the order, the more subsequent volume, with the impacts being stronger in the less liquid deferred live cattle contract.

The impact of transaction size ($TSize_t$) suggests an increase in *ex-post* market relative spread after relatively large transaction sizes. When large orders arrive, the probability of informed trading increases and liquidity providers tend to withdraw their liquidity and set a wider BAS to compensate for higher adverse selection costs (Ho and Stoll 1983; Blume et al. 1994). However, we find none of the $TSize_t$ parameters are statistically significant. The order imbalance ($OIB_{t_{lm}}$) coefficients are small and insignificant in any of the equations.

Consistent with Andersen and Bollerslev (1997) and Wang et al. (2014), our results suggest

persistence of the market quality variables through their lags. In general, larger *ex-ante* trading volume tends to result in higher subsequent market volatility and broader relative spreads. Large *ex-ante* volatility tends to bring more volume and cause wider relative spreads. In both markets and contract series, we see an increase in volatility, volume and spread on announcement days. This effect becomes notably stronger in the 5-minute interval immediately after the announcement releases. This empirical finding is similar to Lehecka et al. (2014), Adjemian and Irwin (2018) and Couleau et al. (2020). Previous studies have shown that time of the day can result in different market conditions. Shang et al. (2018) show that market volatility and BAS are relatively high during the market open, a period when the volume is usually elevated, and the risk to liquidity providers due to adverse selection is high. Similar to findings in the literature, we find increased trading volume, volatility, and relative spread during market opening and closing.

2.6. Discussion and conclusion

Major agricultural commodity exchanges in the U.S. (CME Group and ICE) launched electronic trading platforms by the mid-2000s, resulting in a significant competitive effect on liquidity costs in livestock and grains futures markets (Frank and Garcia 2011 and Wang et al. 2014). However, electronic open limit order books posed new challenges for hedgers, commercials, liquidity providers and other market participants who were slower to adopt their own high-frequency technology, especially those taking large positions in the market. When trading floors were open, these traders could use their discretion to optimally complete their large-sized transactions without price impact (Blume and Goldstein, 1997). However, in electronic open limit order book markets, opportunistic fast traders make order execution more difficult and costly for slow traders. Like other exchanges, CME and ICE offered their solution to this problem by allowing traders to place

hidden limit orders (HLOs) (Frey and Sandas 2017; Degryse et al. 2018).

Our article investigates for the first time the presence of hidden liquidity (HL) in agricultural futures markets and its consequences for market quality. We find evidence that HLOs represent more than 10% (20%) of the total volume in the corn (live cattle) futures market. On average, there exists a higher percentage of HLO in the deferred than the nearby contracts. Exposure of HLOs attracts latent liquidity demand resulting in increased trading volume by 7% (4.4%) in the corn (live cattle) market nearby contract. This occurs as market volatility decreases by 4.4% (6%) and the relative spread diminishes 9.2% (15.4%) in the corn (live cattle) market. Results suggest that increased market quality is related to the buffering effect of HLOs and increased competition for liquidity provision when HLOs are detected. Overall, HLOs reduce volatility, add extra volume and result in lower liquidity costs. These results are robust to different periods and analysis designs. While the classic microstructure theory does not support opacity as it can lead to diminished quality, our results do not support this view. This in turn, offers insights into the degree of success with which traders are hiding and is consistent with iceberg orders being actively used by hedgers, commercials, and institutional investors in commodity and other financial markets (Kaminska 2009; Scaggs 2016). Our results are further supportive of hidden liquidity being especially beneficial in lightly traded markets and may be highly useful in the establishment of new exchanges and markets, particularly if liquidity is poor.

Our empirical findings are also consistent with the view of financial market regulators supporting a certain degree of opacity. While after the financial reform legislation passed in 2010 (“Dodd-Frank Act”), electronic markets have been subject to scrutiny to avoid disrupting trading practices, the Commodity Futures Trading Commission (CFTC 2013) has not limited the use of hidden limit orders. The Security and Exchange Commission (SEC 2016, 2019) has gone a step

further by endorsing non-displayed pegged limit orders and discretionary pegged limit orders that have been adopted by exchanges such as IEX, NYSE American, or Nasdaq to protect investors who want to place hidden passive orders reliably. Pegged limit orders are benchmarked to an index such as the national best bid and offer (NBBO) price and their price is updated in real-time by the exchange, which ensures their placement at the benchmark price. In discretionary orders, discretion is exercised to prevent executions of the orders at unfavorable prices when the market moves against them. These developments suggest the importance of iceberg orders in ensuring that traditional investors have a place in the current fast-trading exchanges.

Despite exchanges and regulators' support for HLOs, fast traders have found their way around opacity by developing their algorithms to detect hidden liquidity and engage in predatory activities that could move the market against HLOs. As these sophisticated algorithms become more efficient, the use of HLOs by traditional market participants may become less attractive. Agricultural commodity exchanges may want to consider alternatives such as pegged discretionary iceberg orders to ensure a proper balance between hedgers and speculators and the prosperity of the futures markets (Working's 1970).

Finally, our study may serve as an input for further research. We know little about the responses of different traders to HLOs. While some traders may speculate for profit when HLOs appear, others may only take the opportunity to build or unfold their positions. Research on these and related questions requires data that clearly identifies HL and traders, which are not currently available. Such information would provide a complete picture to address relevant issues on the market structure design and allow market participants to make more informed choices.

2.7. Tables and Figures

Table 2.1: Summary Statistics

			Volume	Trade Count	Trade Size	BAS	Mean Top Depth	Volatility
Corn	Nearby	Mean	109965	14251	7.716	0.256	221	1.668
		Std. Dev.	43840	5415	1.371	0.016	130	0.992
		Min.	5355	1058	3.342	0.250	7	0.330
		Max.	404100	39212	12.499	0.488	767	7.554
	1st Deferred	Mean	37668	6593	5.713	0.267	97	1.591
		Std. Dev.	40533	5279	1.304	0.014	76	0.938
		Min.	3977	1201	2.579	0.251	11	0.329
		Max.	303568	39431	11.111	0.412	603	6.715
Live Cattle	Nearby	Mean	15612	6374	2.449	0.040	7	0.490
		Std. Dev.	6296	2425	0.270	0.012	13	0.235
		Min.	2328	944	1.780	0.029	3	0.151
		Max.	42227	15346	4.784	0.100	67	1.669
	1st Deferred	Mean	8266	3819	2.164	0.048	4	0.460
		Std. Dev.	4957	2021	0.254	0.014	4	0.237
		Min.	781	378	1.501	0.029	1	0.135
		Max.	27084	13438	2.967	0.109	57	1.712

Notes: The nearby contract is rolled over to the first deferred on the first trading day of the maturity month. All variables are measured exclusively using data of the daytime trading session. Volume is the total count of transacted futures contracts per day. Trade Count measures the number of transactions per day. The Trade Size reports the daily average transaction size in contracts per trade. The BAS is the best ask minus the best bid, calculated each time there is a BAS revision and averaged across revisions within the day. Mean Top Depth is the displayed quoted depth, calculated using the contracts available at the best bid and the best offer as $\frac{BidSize + AskSize}{2}$. The Mean Top Depth is calculated every time there is a revision to the top depth and averaged across all revisions within the day. The Volatility measurement is the standard deviation of the quote midpoints within each day. Summary statistics (Mean, Std. Dev., Min. and Max.) of the distribution across different trading days are offered for each variable.

Table 2.2 Hidden Limit Order (HLO) Ratios

			HLO Volume Ratio	HLO Trade Ratio	Ave. Trade Size (Non-HLO)	Ave. Trade Size (HLO)
Corn	Nearby	Mean	0.113	0.110	5.210	19.745
		Std. Dev.	0.045	0.031	0.813	9.610
		Min.	0.054	0.041	2.703	6.205
		Max.	0.435	0.231	7.687	60.733
	1st	Mean	0.178	0.153	3.780	12.035
	Deferred	Std. Dev.	0.041	0.031	0.700	5.712
		Min.	0.089	0.070	2.309	3.064
		Max.	0.327	0.239	6.548	43.255
Live Cattle	Nearby	Mean	0.231	0.215	2.082	3.807
		Std. Dev.	0.040	0.045	0.213	0.655
		Min.	0.143	0.085	1.640	2.095
		Max.	0.361	0.388	4.265	6.732
	1st	Mean	0.268	0.267	1.818	2.846
	Deferred	Std. Dev.	0.046	0.058	0.228	0.519
		Min.	0.166	0.117	1.318	1.789
		Max.	0.408	0.494	2.987	4.779

Notes: Estimates are daily HLO averages for different contracts and markets. Column 1 reports the daily proportion of trading volume matched against hidden liquidity in corn and live cattle markets for both nearby and first deferred contracts. Column 2 is daily proportion of transactions matched against hidden liquidity. The last two columns report the average transaction sizes that reveal HLO and transaction sizes with no HLO involved. They are calculated every time there is a transaction and averaged across all transactions within the day. Summary statistics (Mean, Std. Dev., Min. and Max.) of the distribution across different trading days are offered for each variable.

Table 2.3 Daily HLO Volume Ratios on USDA Announcement Days

			Non-Announcement HLO Ratio	Announcement HLO Ratio
Corn	Nearby	Mean	0.110	0.136***
		Std. Dev.	0.044	0.041
		Min.	0.054	0.081
		Max.	0.435	0.276
	1st Deferred	Mean	0.175	0.189**
		Std. Dev.	0.044	0.052
		Min.	0.089	0.083
		Max.	0.327	0.288
Live Cattle	Nearby	Mean	0.229	0.276*
		Std. Dev.	0.055	0.083
		Min.	0.143	0.094
		Max.	0.361	0.412
	1st Deferred	Mean	0.267	0.320*
		Std. Dev.	0.063	0.073
		Min.	0.166	0.126
		Max.	0.408	0.391

Notes: This table reports the HLO volume as a percent of total volume on USDA announcement and other trading days. Summary statistics (Mean, Std. Dev., Min. and Max.) of the distribution across different trading days are offered for each variable. The asterisks indicate statistical significance when comparing the mean values between announcement days and other days using the Welch's t-test: *p<0.1; **p<0.05; ***p<0.01.

Table 2.4 Descriptive Statistics for Regression Samples

		Full Sample	Regression Sample	BAS	Mean Trade Size HLO (Non-HLO)	Mean Top Depth	Ann. Days Obs.
Corn	Nearby	4,223,093	69,526	0.256	20.710 (5.007)	227.021	4590 (2828)
	1st Def.	2,000,714	40,522	0.269	11.447 (4.022)	95.779	3076 (1671)
Live	Nearby	2,356,051	43,176	0.044	3.919 (2.737)	7.192	1922
Cattle	1st Def.	1,452,591	22,580	0.047	2.992 (1.813)	3.899	972

Notes: This table presents the descriptive statistics of the selected samples for regression analyses. Full Sample is the total number of transactions recorded during the period of study; Regression Sample is the number of observations selected after the resampling procedure. Within the regression sample, 50% of the transactions involve hidden liquidity (HLO-trade), and the rest have no hidden liquidity involved (non-HLO-trade). The BAS column exhibits the average 1-minute bid-ask spread (BAS) of the regression sample in each market, in cents/bu. (cents/lb.) in corn (live cattle) market. The Mean Trade Size column reports the mean HLO trade size and, in parenthesis, the mean non-HLO trade size. The Mean Top Depth reports the average top-level quoted depth on each trade occurrence. The Ann. Days Obs. column shows the number of observations occurring on USDA announcement days in the regression sample. The numbers in the parentheses are number of observations that occur after the USDA announcements in the corn market.

Table 2.5 Regression Results for the Nearby Contracts

	Corn			Live Cattle		
	Vlt_{t+1}	Vlm_{t+1}	Lqd_{t+1}	Vlt_{t+1}	Vlm_{t+1}	Lqd_{t+1}
BAS_{t-1}	5.121*** (6.621)	-1.226*** (-7.165)		1.407*** (2.915)	-0.407*** (-8.965)	
HLO_t	-0.044*** (-3.655)	0.070** (2.004)	-0.092** (-1.900)	-0.060** (-2.316)	0.044*** (8.952)	-0.154*** (-7.951)
$TSize_t$			1.765 (0.962)			4.042 (0.632)
$OAgr_t$	-1.552* (-1.835)	3.004** (1.952)		-2.371 (-0.855)	8.965 (0.952)	
OIB_{t-1}			-0.002 (-0.652)			0.009 (0.475)
$BfrAnn$	0.446*** (3.542)	0.225 (0.652)	0.473* (1.922)	0.353*** (7.652)	0.106** (2.385)	0.652** (2.244)
$AftAnn_5m$	0.725*** (7.965)	0.662*** (4.968)	0.572*** (3.965)	0.492** (2.226)	0.621*** (4.654)	0.982*** (7.952)
$AftAnn$	0.556** (2.113)	0.358* (1.991)	0.396** (2.196)			
Opn	0.920*** (5.962)	0.717*** (5.851)	0.633* (1.905)	0.772* (1.82)	0.365*** (5.944)	1.135*** (5.342)
Cls	0.439*** (10.695)	0.952 (1.655)	0.634*** (5.547)	1.472*** (6.654)	0.322*** (6.515)	0.964*** (9.432)
Vlt_{t-1}	0.762*** (5.965)	0.636*** (6.581)	0.924*** (5.652)	0.965*** (3.582)	0.452*** (10.855)	0.323*** (8.542)
Vlm_{t-1}	0.534*** (4.751)	0.552*** (3.695)	0.162*** (6.464)	0.415*** (10.965)	0.723*** (10.556)	-1.106 (-0.525)
Lqd_{t-1}			0.908*** (3.005)			0.522*** (7.955)
Constant	0.025 (1.325)	0.455* (1.852)	-0.224** (-2.652)	-1.061 (-0.955)	-0.521*** (-2.932)	0.101 (1.085)
R ²	0.341	0.413	0.399	0.381	0.385	0.339
Observations	69,526	69,526	69,526	43,176	43,176	43,176

Table 2.5 (cont.)

Notes: This table reports the regression coefficients for the corn and live cattle nearby contracts. The t-statistics are reported in parentheses; ***, **, * indicate significance at the 1%, 5% and 10% levels. For each market, the dependent variables are measured for the one-minute interval after a transaction occurs. Vlt_{t+1} , is measured as the coefficient of variation of the quoted midpoints during the 1-minute ex-post interval. Vlm_{t+1} is the 1-minute ex-post volume in number of contracts, normalized by the market's average daily volume for the same month. Lqd_{t+1} is the relative spread measure of the 1-minute interval, calculated as the percentage of quoted bid-ask spread divided by the average dollar value of the top-level quote depth. The smaller the Lqd_{t+1} , the more liquid the market is. All the dependent variables are taken in logarithmic values. BAS_{t-1} is the average bid-ask spread 1 minute before each transaction. HLO_t is a dummy variable that represents whether the transaction that occurs at time t is an HLO-trade (1) or non-HLO-trade (0). $OAgr_t$ measures the aggressiveness of the market order time t, calculated as the prevailing matching side depth on the top level minus the size of the aggressive order. $TSize_t$ is the size of the trade at time t. OIB_{t-1m} is the absolute value of the total order imbalance 1 minute prior to each trade, normalized by the market's average daily volume for the same month. Opn and Cls are variables designed to control for the first 30 minutes of market opening and closing respectively. $BfrAnn$ and $AftAnn$ are dummy variables that measure market changes before and after USDA announcements. $AftAnn_5m$ is a dummy variable that controls for market conditions during the 5 minutes immediately after announcements. In live cattle market, the $AftAnn_5m$ variable measures the first 5-minute of the market open following the after-session announcements on the previous day. Vlt_{t-1} , Vlm_{t-1} and Lqd_{t-1} are variables that measure market volatility, total trading volume, and market liquidity during the 1-minute window prior to the transaction occurrence.

Table 2.6 Regression Results for the Deferred Contracts

	Corn			Live Cattle		
	Vlt_{t+1}	Vlm_{t+1}	Lqd_{t+1}	Vlt_{t+1}	Vlm_{t+1}	Lqd_{t+1}
BAS_{t-1}	2.790*** (4.478)	-1.058*** (-5.824)		1.706*** (4.675)	-0.422*** (-5.652)	
HLO_t	-0.092*** (-2.987)	0.077** (2.362)	-0.012*** (-7.674)	-0.168*** (-5.132)	0.055*** (8.172)	-0.233*** (-4.000)
$TSize_t$			2.010 (0.625)			3.002 (0.444)
$OAgr_t$	-3.155*** (-3.778)	2.226*** (10.002)		-5.009 (-0.167)	4.202** (2.169)	
OIB_{t-1}			0.002 (1.564)			0.001 (0.984)
$BfrAnn$	0.456*** (4.046)	0.247* (1.914)	0.231*** (4.133)	0.726*** (4.711)	0.084*** (3.567)	0.792*** (6.839)
$AftAnn_5m$	0.755*** (3.983)	0.536*** (5.165)	0.545*** (3.946)	0.476** (2.178)	0.566*** (4.646)	1.017*** (8.277)
$AftAnn$	0.560*** (3.878)	0.373*** (5.757)	0.328* (1.952)			
Opn	0.171*** (3.100)	1.731*** (9.927)	1.242*** (11.151)	0.996 (1.666)	1.024*** (8.088)	0.770*** (3.554)
Cls	0.256*** (4.691)	1.518*** (5.857)	0.918*** (3.298)	0.418** (2.253)	0.831*** (10.652)	2.025*** (9.759)
Vlt_{t-1}	0.870*** (10.196)	0.516*** (15.785)	0.853*** (11.375)	0.943*** (7.195)	0.225*** (5.697)	0.759** (2.183)
Vlm_{t-1}	0.372*** (5.640)	0.706*** (10.778)	0.422*** (9.503)	0.658*** (2.791)	0.893*** (15.792)	0.705*** (18.723)
Lqd_{t-1}			0.545*** (21.846)			0.706*** (10.605)
Constant	0.718** (2.138)	0.669 (0.688)	-1.500* (-1.881)	1.098 (0.968)	0.890 (1.144)	-0.696*** (-10.547)
R^2	0.391	0.417	0.310	0.342	0.445	0.339
Observations	40,522	40,522	40,522	22,580	22,580	22,580

Notes: This table reports the regression coefficients for the corn and live cattle first deferred contracts. The t-statistics are reported in parentheses; ***, **, * indicate significance at the 1%, 5% and 10% levels. The variables are defined the same as in table 2.5.

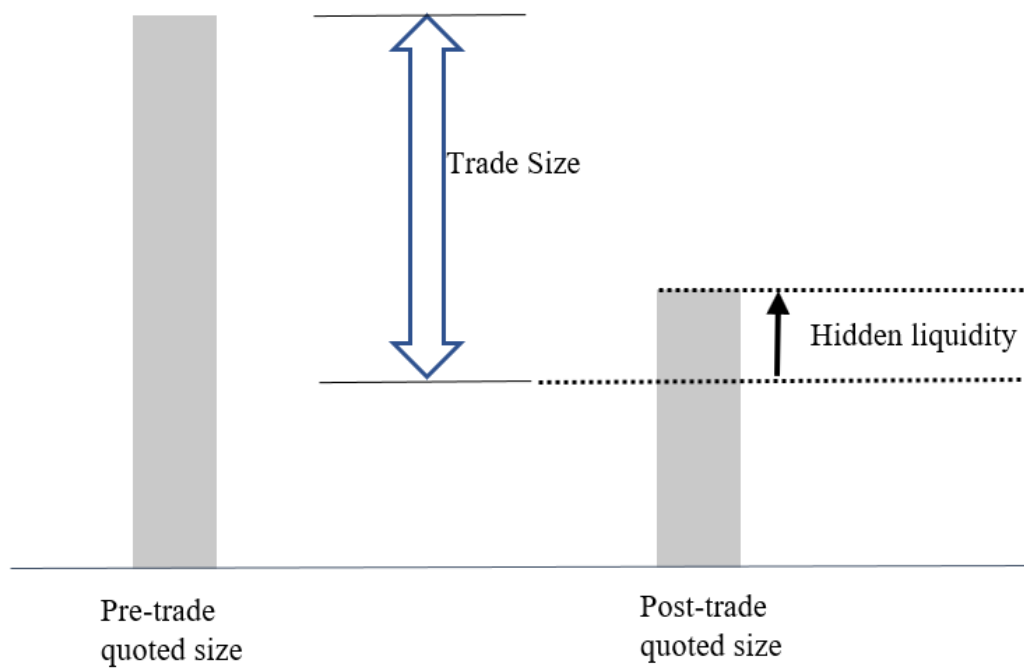


Figure 2.1: Hidden Liquidity in Order Books

Notes: This graph illustrates the underlying reasoning to identify the existence of hidden limit orders by observing the discrepancies between the market order size and the revision of the limit order book.

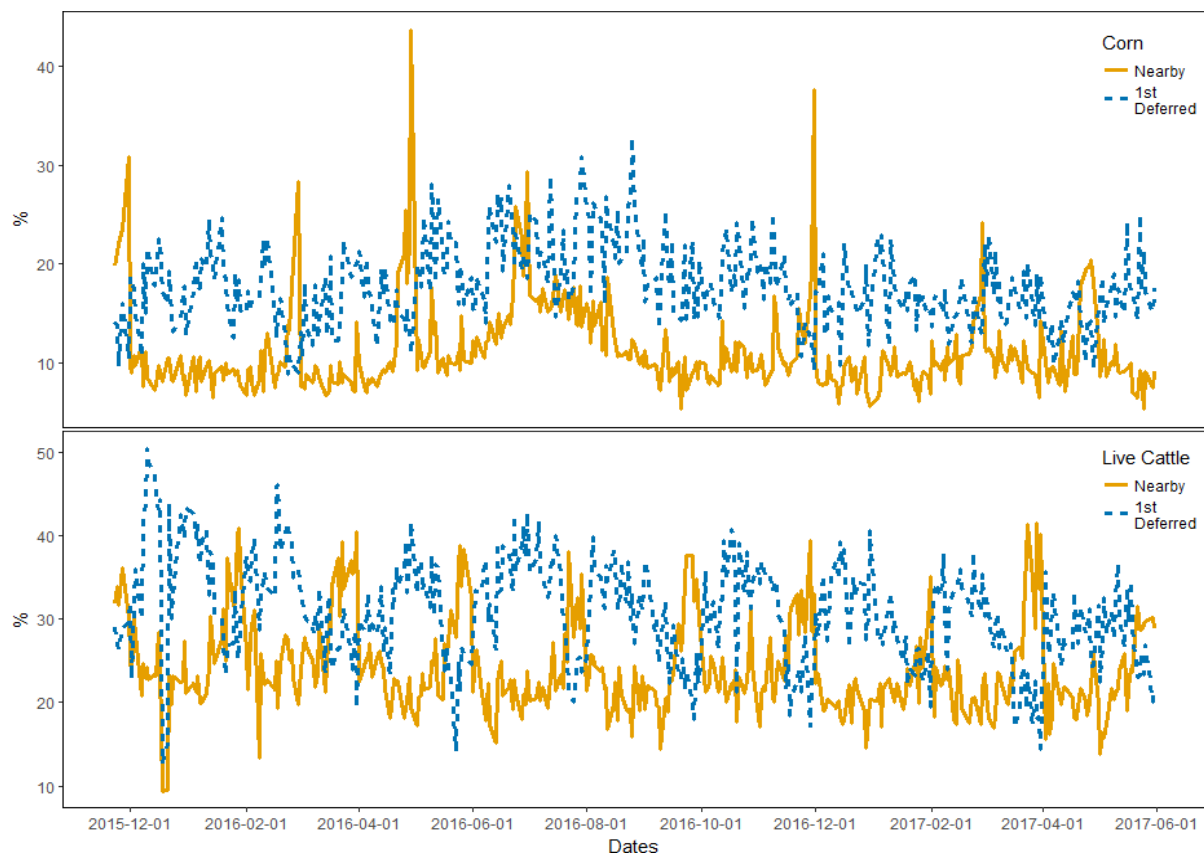


Figure 2.2: Daily Hidden Liquidity Ratios for Corn and Live Cattle Markets (Nearby and First Deferred Contracts)

Note: This figure shows the daily hidden liquidity ratio for the nearby and first deferred contracts in the corn and live cattle futures market. To construct the nearby series, we roll over the nearby contract to the first deferred contract in the first trading day of each contract expiration month. The September corn contract is omitted as discussed in the data section. The top graph shows the hidden liquidity ratio in corn market and the lower graph is for live cattle market. In both plots, solid lines represent the nearby contracts and dashed lines represent the first deferred contracts.

2.8. References

- Adjemian, M. K., & Irwin, S. H. 2018. USDA announcement effects in real-time. *American Journal of Agricultural Economics*, 100(4), 1151-1171.
- Andersen, T. G., & Bollerslev, T. 1997. Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, 4(2-3), 115-158.
- Arzandeh, M., & Frank, J. 2019. Price discovery in agricultural futures markets: should we look beyond the best bid-ask spread? *American Journal of Agricultural Economics*, 101(5), 1482-1498.
- Anand, A., Chakravarty, S., & Martell, T. 2005. Empirical evidence on the evolution of liquidity: Choice of market versus limit orders by informed and uninformed traders. *Journal of Financial Markets*, 8(3), 288-308.
- Bessembinder, H., Panayides, M., & Venkataraman, K. 2009. Hidden liquidity: An analysis of order exposure strategies in electronic stock markets. *Journal of Financial Economics*, 94(3), 361-383.
- Bloomfield, R., O'Hara, M., & Saar, G. 2015. Hidden liquidity: Some new light on dark trading. *The Journal of Finance*, 70(5), 2227-2274.
- Bloomfield, R., & O'Hara, M. 1999. Market transparency: Who wins and who loses? *The Review of Financial Studies*, 12(1), 5-35.
- Blume, L., Easley, D., & O'Hara, M. 1994. Market statistics and technical analysis: The role of volume. *The Journal of Finance*, 49(1), 153-181.
- Blume, M. E., & Goldstein, M. A. 1997: Quotes, order flow, and price discovery. *The Journal of Finance*, 52(1), 221-244.

- Boulatov, A., & George, T. J. 2013. Hidden and displayed liquidity in securities markets with informed liquidity providers. *The Review of Financial Studies*, 26(8), 2096-2137.
- Bryant, H. L., & Haigh, M. S. 2004. Bid–ask spreads in commodity futures markets. *Applied Financial Economics*, 14(13), 923-936.
- Chordia, T., Roll, R., & Subrahmanyam, A. 2001. Market liquidity and trading activity. *The Journal of Finance*, 56(2), 501-530.
- Chicago Mercantile Exchange (CME) Group, *2nd Quarter 2010 Liquidity Monitor*, 2010.
- Available at: https://www.cmegroup.com/education/files/Liquidity_Monitor_070710.pdf
- Commodity Futures Trading Commission (CFTC) 2013. RIN 3038-AD96; Anti-disruptive Practices Authority; Interpretative guidance and policy statement.
- <https://www.cftc.gov/LawRegulation/FederalRegister/FinalRules/2013-12365.html>
- Copeland, T. E., & Galai, D. 1983. Information effects on the bid-ask spread. *The Journal of Finance*, 38(5), 1457-1469.
- Couleau, A., Serra, T. & Garcia, P. 2020. Are corn futures prices getting “jumpy”? *American Journal of Agricultural Economics* 102(2), 569-588.
- Degryse, H. 1999). The total cost of trading Belgian shares: Brussels versus London. *Journal of Banking & Finance*, 23(9), 1331-1355.
- Degryse, H., Karagiannis, N., Tombeur, G. & Wuyts, G. 2018. Two shades of opacity: Hidden orders versus dark trading. <http://dx.doi.org/10.2139/ssrn.2669447>
- De Winne, R., & D'hondt, C. 2007. Hide-and-seek in the market: placing and detecting hidden orders. *Review of Finance*, 11(4), 663-692.
- Easley, D., López de Prado, M. M., & O'Hara, M. 2012. Flow toxicity and liquidity in a high-frequency world. *The Review of Financial Studies*, 25(5), 1457-1493.

- Esser, A., & Mönch, B. 2007. The navigation of an iceberg: The optimal use of hidden orders. *Finance Research Letters*, 4(2), 68-81.
- Frank, J., & Garcia, P. 2011. Bid-ask spreads, volume, and volatility: Evidence from livestock markets. *American Journal of Agricultural Economics*, 93(1), 209-225.
- Frey, S., & Sandås, P. 2017. The impact of iceberg orders in limit order books. *Quarterly Journal of Finance*, 7(03), 1750007.
- Gao, C. 2015. *High frequency trading, hidden orders and market quality in equities*. Doctoral dissertation, Rutgers University-Graduate School-New Brunswick.
- Hasbrouck, J., & Saar, G. 2001. *Limit orders and volatility in a hybrid market: The Island ECN*. Working Paper.
- Hautsch, N., & Huang, R. 2012. *On the dark side of the market: identifying and analyzing hidden order placements*. Available at SSRN 2004231.
- Hayashi, F. 2000. *Econometrics*. Princeton University Press. Section 1: 60-69.
- Haynes, R., & Roberts, J. S. 2015. *Automated trading in futures markets*. White paper, Office of the Chief Economist, Commodity Futures Trading Commission.
- Haynes, R., & Roberts, J. S. 2017. *Automated Trading in Futures Markets-Update*. White paper, Office of the Chief Economist, Commodity Futures Trading Commission.
- Ho, T. S., & Stoll, H. R. 1983. The dynamics of dealer markets under competition. *The Journal of Finance*, 38(4), 1053-1074.
- Hu, Z., Mallory, M., Serra, T., & Garcia, P. 2020 Measuring price discovery between nearby and deferred contracts in storable and non-storable commodity futures markets. *Agricultural Economics*, 51, 825-840.
- Irwin, S. H., & Sanders, D. R. 2012. Financialization and structural change in commodity

- futures markets. *Journal of Agricultural and Applied Economics*, 44(3), 371-396.
- Isengildina, O., Irwin, S. H., & Good, D. L. 2006. The value of USDA situation and outlook information in hog and cattle markets. *Journal of Agricultural and Resource Economics*, 262-282.
- Jain, A., & Jain, C. 2017). Hidden liquidity on the U.S. stock exchanges. *The Journal of Trading* 12 (3), 30-36.
- Joseph, K., & Garcia, P. 2018. Intraday market effects in electronic soybean futures market during non-trading and trading hour announcements. *Applied Economics*, 50(11), 1188-1202.
- Kaminska, I. 2009. *Electronic trading and commodity prices*. *Financial Times*, <https://ftalphaville.ft.com/2009/07/27/63846/electronic-trading-and-commodity-prices/>
- Krinsky, I., & Lee, J. 1996). Earnings announcements and the components of the bid-ask spread. *The Journal of Finance*, 51(4), 1523-1535.
- Kovaleva, P., & Iori, G. 2015. The impact of reduced pre-trade transparency regimes on market quality. *Journal of Economic Dynamics and Control*, 57, 145-162.
- Lehecka, G. V., Wang, X., & Garcia, P. 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.
- Moinas, S. 2010). *Hidden limit orders and liquidity in order driven markets*. Working Paper
- Newey, W. K., & West, K. D. 1987. Hypothesis testing with efficient method of moments estimation. *International Economic Review*, 28(3), 777-787.
- Pardo, A., & Pascual, R. 2012. On the hidden side of liquidity. *The European Journal of Finance*, 18(10), 949-967.

- Roşu, I. 2019. Fast and slow informed trading. *Journal of Financial Markets*, 43, 1-30.
- Scaggs, A. 2017. *NYSE changes its mind about favouring dark liquidity*. Financial Times.
<https://ftalphaville.ft.com/2017/01/25/2183155/nyse-changes-its-mind-about-favouring-dark-liquidity/>
- Securities and Exchange Commission (SEC) 2016. Release No. 34-78101; File No. 10-222; June 17, 2016; In the Matter of the Application of Investors' Exchange, LLC for Registration as a National Securities Exchange. <https://www.sec.gov/rules/other/2016/34-78101.pdf>
- Securities and Exchange Commission (SEC) 2019. Release No. 34-85158; File No. SR-NYSE-2018-52; February 15, 2019; Self-Regulatory Organizations; New York Stock Exchange LLC; Order Granting Approval of a Proposed Rule Change to Amend NYSE Rule 7.31.
<https://www.sec.gov/rules/sro/nyse/2019/34-85158.pdf>
- Shang, Q., Mallory, M., & Garcia, P. 2018. The components of the bid-ask spread: Evidence from the corn futures market. *Agricultural Economics*, 49(3), 381-393.
- Stoll, H. R. 1978. The supply of dealer services in securities markets. *The Journal of Finance*, 33(4), 1133-1151.
- Tuttle, L. 2006. *Hidden Orders, Trading Costs and Information*. Working Paper.
- U.S. Department of Agriculture. 2012 *Public Comments Received in Response to Federal Register Notice*, June 8, 2012. Available at:
https://www.nass.usda.gov/Newsroom/2012/Public_Comments_Release_Times.pdf
- Wang, X., Garcia, P., & Irwin, S. H. 2014. The behavior of bid-ask spreads in the electronically-traded corn futures market. *American Journal of Agricultural Economics*, 96(2), 557-577.
- Wang, G. H., & Yau, J. 2000. Trading volume, bid-ask spread, and price volatility in futures

markets. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, 20(10), 943-970.

Working, H. 1970. *Economic Functions of Futures Markets*, " a chapter in *Futures Trading in Livestock—Origins and Concepts* (Edited by H. Bakken). Chicago Mercantile Exchange.

2.9. Supplementary Results

2.9.1. *Description of CME Group Market Depth (MD) data*

The MD data are time-stamped to the nearest millisecond and contain market data messages necessary to recreate the limit order book (LOB) up to ten levels deep in the corn market and five levels in the live cattle market. Trading times and contract specifications differ in the two markets. Trading in the corn futures market occurs in the evening session from 7:00PM to 7:45AM (on the next day) and the regular day trading session from 8:30 AM to 1:20 PM. The minimum price movement (tick size) in corn futures market is 0.25 cents/bu. Corn futures contracts have five delivery months: March, May, July, September, and December. Since February 28, 2016, when trading hours were changed, the live cattle trading hours are from 8:30AM to 1:05PM. Before the change, the hours varied across different weekdays. From November 23, 2015 to February 27, 2016, live cattle trading hours were 9:05AM to 4:00PM on Monday; 8:00AM to 4:00PM from Tuesday to Thursday; and 8:00AM to 1:55PM on Friday. Evening trading in the live cattle market was discontinued on October 27, 2014. The available delivery months for live cattle contracts are February, April, June, August, October, and December. The tick size is 0.025 cents/lb. in the live cattle futures market. We include only day trading hours for the corn market, where most of the trading activity occurs.

2.9.2. *HLO detection algorithm*

Table 2.7 in the supplementary results demonstrates how the HLO detection algorithm works by presenting how market orders matched against HLO can be identified using the Market Depth dataset. Table 2.7 represents the public LOB and thus does not display hidden liquidity. We classify market orders in three groups, depending on their size relative to the resting limit orders

awaiting execution in the LOB. We assume the incoming market order is a buy order for 50 contracts.

In case 1, we assume that there are hidden limit orders lying within the spread willing to sell at \$96 per contract. We now demonstrate how the algorithm can produce a conservative estimate of the number of contracts hidden. Table 2.7 in the supplementary results shows that the incoming buy market order is fully matched with the hidden liquidity and executed within the quoted spread at \$96. Notice that this causes no change in the quotes resting in the LOB in supplemental table 2.7. Also, notice that the HL identified is for 50 contracts. Any hidden quantity higher than 50 contracts will remain undetected.

Case 2 shows the market buy order being matched to the best ask price displayed in the book. Here, HL is assumed to be lying at the top of the book and thus to contain sell orders for \$96. In the absence of HL, a 50-contract market order would be filled as follows: 26 contracts bought at \$96 and the remaining 24 contracts bought at \$97. However, we observe the after-trade ask price at \$96 with 13 available contracts. The difference is due to detected HL which equals 37 contracts ($50 \text{ (buy market order)} - 26 \text{ (best ask depth before buy market order execution)} + 13 \text{ (best ask depth after buy market order execution)} = 37$).

In case 3 the market order exhausts the entire depth at the first level and moves up the book to fulfill the order. HLOs are assumed both at the top of the book containing sell orders at \$96 and then at the second level of the book, containing sell orders at \$97. The buy market order is executed as follows: 38 contracts are purchased at \$96 and 12 at \$97. At \$96, the detected hidden order size equals 13 contracts ($38 \text{ (buy market order executed at \$96)} - 25 \text{ (best ask at \$96 before market execution)} = 13$). There are 2 additional contracts revealed from HL at \$97 ($12 \text{ (buy market order$

executed at \$97) - 80 (best ask at \$97 before market execution) +70 (best ask at \$97 buy market order execution) = 2).

2.9.3. HLO Detection Example

Table 2.7: HLO After a Buy Order of 50 Contracts

Displayed LOB Before Trade		Trade Occurrence		Displayed LOB After Trade	
	Price (USD/contract)	Depth (contract)		Price (USD/contract)	Depth (contract)
Case #1					
Asks	100	10	Trade Price: 96	100	10
	99	75	Trade Size: 50	99	75
	98	23	Detected HLO Price: 96	98	23
Bids	95	17	Detected HLO Size: 50	95	17
	94	32		94	32
	93	45		93	45
Case #2					
Asks	100	10	Trade Price: 96	100	10
	97	75	Trade Size: 50	97	75
	96	26	Detected HLO Price: 96	96	13
Bids	95	17	Detected HLO Size: 50-26+13=37	95	17
	94	32		94	32
	93	45		93	45
Case #3					
Asks	100	10	Trade Price1: 96	101	32
	97	80	Trade Size1: 38	100	10
	96	25	Detected HLO Price1: 96	97	70
Bids	95	17	Detected HLO Size1: 38-25=13	95	17
	94	32	Trade Price2: 97	94	32
	93	45	Trade Size2: 12 Detected HLO Price2: 97 12-80+70=2	93	45

Notes: This table presents a limit order book under three scenarios in which hidden liquidity can be identified. The ask levels are boldfaced. The first two columns show the quoted price (column 1) and the corresponding quoted depth (column 2) before a transaction occurs. The last two columns show the quoted price and the corresponding quoted depth after a transaction occurs. We assume a buy market order with size 50 contracts arrives. Hidden liquidity for sell market orders can be identified in the same fashion.

2.9.4. HLO Volume Cases

Table 2.8: HLO Volume by Case

			Case 1 Volume	Case 2 Volume	Case 3 Volume
Corn	Nearby	Mean	586	7912	3200
		Std. Dev.	526	3457	2581
		Min.	16	108	403
		Max.	3872	33085	22110
	1st Deferred	Mean	726	3483	2193
		Std. Dev.	534	3783	2187
		Min.	74	263	70
		Max.	3537	24885	22242
Live Cattle	Nearby	Mean	597	714	1049
		Std. Dev.	271	357	486
		Min.	113	45	118
		Max.	1992	2950	2950
	1st Deferred	Mean	557	284	582
		Std. Dev.	245	284	409
		Min.	63	0	36
		Max.	1561	1603	2545

Notes: Columns 1 to 3 report daily HLO averages of the transaction volume for three different scenarios discussed in supplemental table 2.7 Case 2 is the most common channel through which HLO reveals itself in the corn market, while case 3 is most common in the live cattle market. This suggests a vast majority of corn market orders are filled at the best quoted level. In contrast, in the live cattle market, results suggest that many orders cannot be filled at the first level of market depth. This is consistent with findings in previous studies that BAS is relatively larger in the live cattle market due to thinner market depth (e.g. Frank and Garcia 2010; Arzandeh and Frank 2019).

2.9.5. Robustness Checks

We examine the robustness of the main results by considering different *ex-ante* and *ex-post* trade intervals and by increasing the transactions not revealing HLOs (non-HLOs). Changing the intervals and increasing non-HLOs allows us to assess the intensity of the HLO effect at shorter time periods and the representativeness of our findings. We conduct the analyses using observations from the nearby contracts in 2016 which was a volatile year especially for corn prices. Heightened volatility and use of nearby contracts will ensure adequate HLO activity.

The nearby contract regression results are in the supplementary results (table 2.9). These results are a replication of the results in the main document for the specific year (2016) and contract (nearby) chosen. Overall, results are similar to those presented in the manuscript. The HLOs contribute to improved market quality by reducing the relative spread and volatility, while increasing trading volume. Notice in this highly volatile period, HLOs seem to be more effective in reducing volatility and improving trade, but less effective in reducing the relative spread.

In supplemental table 2.10. we present the results of changing the duration of the *ex-ante* and *ex-post* time intervals from 1-minute to 30-seconds. Since overlapping is now less frequent, the number of observations increase. Overall, we see slightly stronger impacts of the HLOs on market quality when shorter time intervals are used. This finding is consistent with Pardo and Pascual (2012), who conclude that the impacts of HLOs are relatively short-lived.

In the last robustness analysis, we change the structure of the sample to include more non-HLOs observations. The observations are not matched; all HLOs and all non-HLO observations are included provided the observations do not overlap using our *ex-ante* and *ex-post* intervals. Results are in supplemental table 2.11 and 2.12 for the 1-minute and 30-second intervals, respectively. Relaxing the pairwise matching increases the sample size significantly. Overall the

results do not change using the 1-minute interval and as seen earlier become stronger using the 30-second interval. On balance, the impacts of HLO are robust to different periods and empirical analysis designs.

2.9.6. Regression Results for the Nearby Contracts, 2016 (1-min interval)

Table 2.9: Regression Results for the Nearby Contracts, 2016 (1-min interval)

	Corn			Live Cattle		
	Vlt_{t+1}	Vlm_{t+1}	Lqd_{t+1}	Vlt_{t+1}	Vlm_{t+1}	Lqd_{t+1}
BAS_{t-1}	4.858*** (3.161)	-0.408*** (-5.409)		0.523 (0.508)	-0.100*** (-8.784)	
HLO_t	-0.047* (-2.026)	0.081*** (3.166)	-0.068 (-2.012)	-0.139** (-2.479)	0.098*** (4.657)	-0.097*** (-3.629)
$TSize_t$			2.814 (1.234)			2.643 (1.214)
$OAgr_t$	-2.378 (-0.935)	2.579 (1.221)		-8.847 (-0.656)	6.524 (1.327)	
OIB_{t-1}			0.005 (1.035)			0.006* (1.876)
$BfrAnn$	0.749 (1.015)	1.630** (2.656)	0.829*** (3.432)	0.167*** (4.377)	0.655 (1.055)	1.477* (1.910)
$AftAnn_{5m}$	1.014*** (4.179)	0.883*** (5.009)	0.851*** (4.064)	1.451** (2.140)	1.443*** (4.396)	1.061*** (7.957)
$AftAnn$	1.081** (2.425)	1.722* (1.932)	1.007*** (3.832)			
Opn	0.242*** (3.526)	0.240*** (3.789)	0.752* (1.943)	5.576*** (8.923)	-0.631 (1.593)	0.124*** (5.624)
Cls	0.982* (1.836)	0.384*** (6.235)	1.479 (0.993)	0.481*** (11.062)	0.590*** (4.500)	0.473*** (8.076)
Vlt_{t-1}	0.600*** (11.053)	0.161*** (12.516)	0.567*** (10.983)	1.489*** (14.330)	0.205*** (3.417)	0.080*** (11.696)
Vlm_{t-1}	0.114*** (4.312)	0.904*** (11.847)	0.579*** (5.418)	0.986*** (4.251)	0.708*** (12.093)	0.572 (0.655)
Lqd_{t-1}			0.512*** (10.416)			0.427*** (5.821)
Constant	-0.624 (-0.505)	1.729 (1.110)	-2.411 (1.085)	3.303 (4.731)	0.510 (7.517)	-0.547 (-1.464)
R^2	0.311	0.386	0.412	0.390	0.395	0.347
Observations	48,668	48,668	48,668	28,496	28,496	28,496

Notes: The observations are matched (1 HLO and 1 non-HLO) and separated by at least a 1-minute *ex-ante* and *ex-post* interval. Details on the matching algorithm are in the text. Variable definitions are in the text and corresponding tables in the article.

2.9.7. Regression Results for the Nearby Contracts, 2016 (30-sec interval)

Table 2.10: Regression Results for the Nearby Contracts, 2016 (30-sec interval)

	Corn			Live Cattle		
	Vlt_{t+30}	Vlm_{t+30}	Lqd_{t+30}	Vlt_{t+30}	Vlm_{t+30}	Lqd_{t+30}
BAS_{t-30}	0.481 (1.507)	-1.035*** (-10.773)		1.016*** (15.030)	-1.299*** (-12.108)	
HLO_t	-0.170*** (-3.572)	0.245*** (7.157)	-0.080*** (-3.153)	-0.110*** (-6.258)	0.118*** (6.125)	-0.042** (-2.193)
$TSize_t$			1.567 (0.359)			0.731 (0.017)
$OAgr_t$	-1.321 (-0.235)	1.801*** (3.230)		-7.891 (-0.428)	2.306 (0.974)	
OIB_{t-30}			1.465 (0.324)			4.965 (1.015)
$BfrAnn$	0.525 (0.878)	0.705 (1.147)	0.608*** (3.451)	1.308 (0.980)	0.905 (0.918)	1.329* (1.899)
$AftAnn_{5m}$	1.725*** (4.133)	1.309*** (4.764)	0.800*** (3.602)	1.409** (2.253)	1.233*** (4.416)	1.202*** (7.787)
$AftAnn$	0.589*** (4.774)	1.466*** (6.007)	1.034** (2.292)			
Opn	3.641*** (4.122)	2.045*** (6.913)	1.675*** (4.253)	2.379*** (8.797)	1.997*** (8.768)	1.933*** (4.030)
Cls	0.676*** (10.505)	0.467*** (6.741)	0.037 (1.764)	1.226 (1.283)	0.443 (1.254)	0.184*** (5.852)
Vlt_{t-30}	0.381*** (3.669)	0.035*** (9.251)	0.095*** (7.963)	0.712*** (4.839)	0.041*** (15.181)	0.138*** (2.854)
Vlm_{t-30}	0.341*** (2.986)	0.120*** (15.863)	-0.273 (-1.219)	0.717*** (5.164)	0.454*** (14.351)	0.360 (1.137)
Lqd_{t-30}			0.675*** (3.077)			0.617*** (8.208)
Constant	-1.033*** (-3.337)	-0.042 (-0.418)	-0.631*** (-5.799)	2.872*** (5.446)	0.794*** (9.747)	2.169*** (15.220)
R^2	0.371	0.318	0.279	0.246	0.311	0.395
Observations	169,146	169,146	169,146	92,359	92,359	92,359

Notes: The observations are matched (1 HLO and 1 non-HLO) and separated by at least a 30-second *ex-ante* and *ex-post* interval. Details on the matching algorithm in the text. Variable definitions are in the text and corresponding tables in the article.

2.9.8. Regression Results for the Nearby Contracts, 2016 (1-minute interval, no match)

Table 2.11: Regression for the Nearby Contracts, 2016 (1-minute interval, no match)

	Corn			Live Cattle		
	Vlt_{t+1}	Vlm_{t+1}	Lqd_{t+1}	Vlt_{t+1}	Vlm_{t+1}	Lqd_{t+1}
BAS_{t-1}	1.966*** (2.981)	-1.757*** (-4.288)		2.990*** (6.749)	-0.490*** (-6.901)	
HLO_t	-0.117*** (-3.246)	0.053** (2.489)	-0.145*** (-4.064)	-0.081 (-0.537)	0.048*** (3.238)	-0.155*** (-4.625)
$TSize_t$			1.073 (0.070)			1.577 (0.049)
$OAgr_t$	-2.216 (-0.417)	2.029 (0.602)		-0.526 (-0.015)	3.152 (0.231)	
OIB_{t-1}			0.249 (0.041)			-0.900 (-0.265)
$BfrAnn$	0.436 (0.619)	0.478 (0.997)	0.531 (1.217)	1.751* (2.010)	0.741 (1.291)	0.929 (1.554)
$AftAnn_5m$	1.430*** (3.591)	1.221*** (5.296)	0.671*** (4.092)	2.440* (2.030)	1.545*** (4.397)	1.011*** (7.545)
$AftAnn$	0.942*** (3.102)	0.530*** (5.344)	0.492*** (2.736)			
Opn	2.418*** (11.392)	2.348*** (9.938)	1.885*** (11.465)	2.243*** (7.872)	1.786*** (7.394)	1.768*** (7.766)
Cls	0.732*** (10.248)	0.434*** (14.522)	1.012*** (11.625)	1.484*** (7.047)	0.191*** (5.198)	0.232*** (5.745)
Vlt_{t-1}	0.835*** (6.720)	0.024*** (3.748)	0.059*** (7.629)	1.138*** (9.672)	0.067*** (9.401)	0.058*** (8.057)
Vlm_{t-1}	0.480*** (16.329)	0.126*** (6.341)	0.018 (0.907)	0.709*** (7.626)	0.106*** (6.672)	-0.155 (-0.940)
Lqd_{t-1}			0.944*** (3.940)			0.682*** (4.560)
Constant	-0.496 (-0.588)	0.890* (1.937)	-0.048 (-0.194)	-1.332 (-1.245)	-0.531*** (-4.379)	-0.601*** (-3.824)
R^2	0.301	0.308	0.311	0.242	0.299	0.350
Observations	111,550	111,550	111,550	58,900	58,900	58,900

Notes: The observations are not matched; all HLOs and all non-HLO observations are included provided the observations do not overlap using a 1-minute *ex-ante* and *ex-post* interval. Variable definitions are in the text and the corresponding tables in the article.

2.9.9. Regression Results for the Nearby Contracts, 2016 (30-second interval, no match)

Table 2.12: Nearby Contracts, 2016 (30-second interval, no match)

	Corn			Live Cattle		
	Vlt_{t+30}	Vlm_{t+30}	Lqd_{t+30}	Vlt_{t+30}	Vlm_{t+30}	Lqd_{t+30}
BAS_{t-30}	1.811** (2.514)	-1.216*** (-5.984)		0.565*** (10.354)	-1.584*** (-9.824)	
HLO_t	-0.199*** (-3.946)	0.265*** (8.141)	-0.323*** (-11.119)	-0.123** (-2.239)	0.150*** (3.540)	-0.265*** (-7.478)
$TSize_t$			1.362 (0.083)			0.968 (0.035)
$OAgr_t$	-3.638 (-0.753)	2.110 (0.564)		-10.952 (-0.542)	2.446 (0.134)	
OIB_{t-30}			0.864 (0.203)			5.424 (1.084)
$BfrAnn$	0.735* (2.035)	0.417 (1.675)	0.407* (1.903)	2.705*** (16.892)	1.010*** (16.450)	0.506*** (16.087)
$AftAnn_{5m}$	1.610*** (3.597)	1.830*** (4.667)	0.914*** (3.688)	1.447* (1.974)	1.536*** (4.860)	0.951*** (8.230)
$AftAnn$	1.368*** (4.549)	1.105*** (7.028)	0.592*** (5.694)			
Opn	3.075*** (5.284)	2.608*** (5.350)	1.839*** (4.227)	3.068*** (17.399)	2.082*** (15.782)	2.614*** (16.759)
Cls	1.428*** (15.436)	0.697*** (14.883)	0.130*** (15.891)	1.809*** (15.930)	0.451*** (9.482)	0.465*** (21.485)
Vlt_{t-30}	0.915*** (8.290)	0.056*** (6.469)	0.011*** (4.322)	0.672*** (7.861)	0.059*** (14.176)	0.035*** (9.873)
Vlm_{t-30}	0.646*** (31.020)	0.124*** (12.163)	0.036*** (3.904)	0.504*** (18.315)	0.385*** (20.503)	0.087*** (7.579)
Lqd_{t-30}			0.921*** (2.854)			0.676*** (13.933)
Constant	-0.910 (-1.155)	-0.103 (-0.481)	-0.442*** (-2.840)	-1.180*** (-6.846)	0.093 (-0.507)	0.895*** (-6.770)
R^2	0.296	0.315	0.313	0.233	0.299	0.353
Observations	524,355	524,355	524,355	255,835	255,835	255,835

Notes: The observations are not matched; all HLOs and all non-HLO observations are included provided the observations do not overlap using a 30-second *ex-ante* and *ex-post* interval. Variable definitions are in the text and the corresponding tables in the article.

CHAPTER 3

RIDE THE TREND: IS SPREAD MOMENTUM A RELIABLE TRADING INDICATOR IN THE U.S. COMMODITY MARKETS?

3.1. Introduction

In an efficient market, the current price reflects all available information, so future price changes are unpredictable based on this information. Time series momentum (TSM) is a market pricing anomaly where returns are characterized by apparently persistent trends that are taken to predict future returns and offer profit opportunities. This phenomenon, studied initially by Jegadeesh and Titman (1993), has been the focus of multiple studies covering a wide array of financial markets (e.g. Miffre and Rallis 2007; Moskowitz et al. 2012; Asness et al. 2013; Daniel and Moskowitz 2016; Georgopoulou and Wang 2017; Wang et al. 2020). Researchers have generated abnormal positive returns from TSM investment strategies by applying sophisticated momentum filtering and portfolio rebalancing techniques. Persistent TSM is a major challenge to market efficiency, since it implies that investors can consistently profit from market imperfections by chasing price trends. Widespread acceptance of the efficient market hypothesis (Fama 1970) has led many academics to express skepticism about profitability of TSM-based investment strategies (e.g. Park and Irwin 2007; Huang et al. 2020).

Research reporting the existence of TSM can be questioned because of weaknesses in the empirical design and statistical procedures. The main empirical design issue is the inappropriate treatment of futures prices during the rollover months. Futures contracts are legal instruments involving two parties, one interested in buying (long position) and the other interested in selling (short position) the contract's underlying asset at a certain price and time in the future. In contrast to stocks and other financial instruments, each futures contract has an expiration date. In order to

hold positions in the market beyond that date, investors usually close positions in the contract approaching expiration (nearby contract) and build new positions in a deferred contract. This has implications for return calculation: futures investors' returns are based on the contract's price change from the time of purchase to sale. For example, assume an investor who takes a long outright position in the corn futures market by buying one September contract for \$25,500 and selling it before expiration for \$25,700. The difference between the sale and purchase price turns into a \$200 profit. Subsequently, the investor buys a December contract for \$26,000 and sells it later for \$26,200, which turns into another profit of \$200. The total gain from the two contracts is \$400. Several studies have created the prospect of equity-like returns in futures markets (e.g. Miffre and Rallis 2007, Moskowitz et al. 2012; Boons and Prado 2019) by including the non-realizable roll yield. In our example, this is equivalent to calculating returns from the two transactions as the difference between the December contract sale price (\$26,200) and the September contract purchase price (\$25,500). This results in a profit of \$700 which includes the realizable profit of \$400 and the roll yield of \$300. The roll yield is the difference between the purchase price of the December contract (\$26,000) and the sale price of the September contract (\$25,700). Since rolling from one contract to another involves no cash flows to investors, calculating futures returns by including the contract roll yield is inappropriate and misleading (Singleton, 2014; Bessembinder 2018; Irwin et al. 2020; Paschke et al. 2020).

Another design issue affecting previous momentum studies concerns volatility scaling. In particular, researchers have built portfolios that take greater (smaller) positions in assets with lower (larger) volatility to ensure their returns are not dominated by high volatility assets. Kim et al. (2016) show that this practice leads to confounding TSM with volatility-scaling returns. By

properly allowing for contract rollover, Kim et al. (2016) find that TSM profits in Moskowitz et al. (2012) are mainly driven by asset allocation rather than TSM.

A necessary condition for TSM to be a profitable investment strategy is that past returns predict future returns. Return predictability has been typically assessed through a regression analysis that suffers from several statistical issues. Returns across different assets are usually pooled in a regression analysis that does not control for differences in average returns across markets. This tends to generate upward bias in the estimates. This upward bias is exacerbated by the usual practice of scaling pooled returns by volatility in the regression analysis. The use of a predictor that measures the return during several past months can result in further upward bias and over-rejection of the null hypothesis of no predictability due to persistence (Ang and Bekaert 2007; Huang et al. 2020). After correcting for different return means and variances across markets, bootstrap methods are required to derive proper critical values for the predictability test (Huang et al. 2020). These authors conclude that evidence of time series momentum is virtually nil, based on proper specification of predictability regressions and tests.

Recent studies have revisited TSM by addressing only a limited number of the multiple issues affecting previous research, thus providing an incomplete assessment of the feasibility of momentum returns. Also, while the presence of TSM in outright futures markets has been widely investigated, recent research has provided evidence of profitability related to calendar spread time series momentum (STSM), both in individual commodity markets and portfolios (e.g. Szymanowska et al. 2014; Boons and Prado 2019). While an outright position implies that the trader takes only one position in the market (long or short), a calendar-spread investment strategy involves simultaneously taking two offsetting positions on the same underlying asset using contracts with different expiration dates. For example, the investor may hold two positions in the

corn futures market at the same time, as the seller of a July contract and as the buyer of a December contract. The STSM literature suffers from the same empirical and statistical issues discussed earlier. Unlike previous studies, we assess the existence of STSM and the profitability of STSM investment strategies by properly addressing previous research issues.

Boons and Prado (2019) suggest that STSM not only predicts future spread returns, but also offers high predictive power of outright returns. Other studies find the roll yield magnitude to predict future outright returns (e.g. Gorton et al. 2013, Irwin et al. 2020), with higher (lower) roll yield indicating higher (lower) future outright returns. Consistent with their findings, we investigate the existence of STSM-spread and STSM-outright. We consider a wide array of US commodity futures markets, which are similar to the markets considered in Szymanowska et al. (2014) and Boons and Prado (2019), the two most influential studies on STSM. First, we assess whether past spread returns can predict future spread returns (STSM-spread) and outright returns (STSM-outright). Second, we study the profitability of STSM-spread and STSM-outright investment strategies.

Szymanowska et al. (2014) sort 21 assets into four quartiles based on their market momentum. They conclude that markets in the highest momentum quartile tend to generate larger spread returns compared to those with lowest momentum, and this momentum strategy is robust to different holding periods. Boons and Prado (2019) investigate the STSM returns through a multistep sorting process and build a portfolio of eight assets: the four assets with the highest and the lowest spread momentum. They provide evidence that an increase in STSM results in larger spread and outright portfolio returns. Unlike these studies, which provide evidence of momentum based on selective and restrictive portfolios, we take a more direct and comprehensive approach by considering a portfolio that includes all markets. This facilitates replicability and reduces the

risk that selective portfolio strategies favor markets with high volatility and large execution costs that can erode returns (Lesmond et al. 2004).

To examine STSM-spread and STSM-outright returns, we compare them against returns derived from a benchmark, defined as the static long-spread and long-outright investment strategy for the STSM-spread and STSM-outright, respectively. Compared with the static long-only strategy, momentum is an active strategy that requires regularly updating the momentum indicator and consequent reassessment of the investor position. Strong momentum should result in a genuine advantage of the STSM over the long-only strategy.

Our research is timely as U.S. commodity futures markets have recently experienced a major transformation that may have changed price behaviour and affected predictability of the futures term structure. The transition from pit to electronic trading over the last decades has facilitated market access and resulted in the financialization of commodity markets (Irwin and Sanders 2011, Sanders and Irwin 2012). Our research results are relevant to investors, hedgers, commercials and regulators. Kang et al. (2020) have demonstrated that in addition to the passive long-only strategies traditionally adopted by financial funds (Irwin et al. 2020 and Moran et al. 2020), they also use momentum strategies. Our comparison between the two strategies is of particular interest to financial fund managers. Existence of TSM would challenge market efficiency and the market's ability to perform its functions of facilitating risk management and price discovery. Futures market inefficiencies may reduce correlation between cash and futures prices, increase basis risk⁷ and the efficacy of futures contracts as risk management tools. Hence, our research is also relevant to those

⁷ The basis is the difference between the futures' contract underlying asset cash price and the futures price.

who use futures markets to hedge and inform their risk management and inventory decisions (Hieronymus 1977; Tomek 1997). Finally, our analysis is relevant to regulators with a mission to promote market efficiency.

3.2. Trading strategies

Implementing a time series momentum trading strategy involves a look-back period (L) and a holding period (H). The look-back period L is the reference period to calculate the momentum indicator used to establish a market position during the holding period H. A positive (negative) momentum indicator occurs when the market displays a positive (negative) cumulative return during the look-back period. When the momentum indicator is positive (negative), a long (short) position is taken in the next holding period and maintained for a period H. The market momentum is re-evaluated after each holding period and the market position is unchanged or reversed, depending on the updated momentum indicator. Most previous studies calculate the market momentum using a 12-month look-back period and 1-month holding period (Moskowitz et al. 2012; Asness et al. 2013; Boons and Prado 2019; Huang et al. 2020). While we follow this convention, we also provide results with alternative look-back periods.

Unlike equity markets, constructing a time series of monthly “nearby contract” prices in a futures market involves using contracts with different maturities, which needs to be considered when computing monthly returns. We compute returns as in equation (3.1), which ensures they are not calculated across contracts and thus exclude the roll yields (Singleton 2014; Paschke et al. 2020):

$$R_{t+1}^{nearby} = \begin{cases} \frac{p_{t+1}^{nearby}}{p_t^{nearby}} - 1 & \text{if no contract rollover} \\ \frac{p_{t+1}^{Tn}}{p_t^{Tn}} - 1 & \text{if contract rollover} \end{cases} . \quad (3.1)$$

When no rollover occurs between month t and $t+1$, returns are calculated as $\frac{P_{t+1}^{nearby}}{P_t^{nearby}} - 1$, where P_{t+1}^{nearby} (P_t^{nearby}) is the closing price of the nearby contract at the end of month $t+1$ (t). When there is contract rollover, returns are derived as $\frac{P_{t+1}^{T_n}}{P_t^{T_n}} - 1$ where T_n is the new nearby contract (first deferred in the previous month t), and $P_{t+1}^{T_n}$ ($P_t^{T_n}$) represents the closing price at the end of month $t+1$ (t).

Following the momentum literature, we define the nearby contract momentum indicator (M) as:

$$M_t^{nearby} = \prod_{s=t+1-L}^t (1 + R_s^{nearby}) - 1, \quad (3.2)$$

with the first deferred momentum indicator computed in a similar way. Notice monthly returns accumulate over the look-back period. The spread momentum (SM) indicator is calculated using the nearby minus the first deferred contract momentum (Szymanowska et al. 2014; Boons and Prado 2019):

$$SM_t = \prod_{s=t+1-L}^t (1 + R_s^{nearby}) - \prod_{s=t+1-L}^t (1 + R_s^{deferred}). \quad (3.3)$$

When the nearby contract price increases by more (less) than the deferred contract during the look-back period, the spread carries a positive (negative) momentum, i.e. $SM_t > 0$ ($SM_t < 0$). A positive (negative) momentum could also occur when the nearby contract price declines less (more) than the deferred contract. A STSM investment strategy establishes a long (short) spread/outright position when $SM_t > 0$ ($SM_t < 0$). The long (short) spread position consists of a long (short) position in the nearby and a short (long) position in the deferred contract. Spread and outright momentum returns for the holding period $H=1$ are:

$$R_{t+1}^{spread} = \text{sign}(SM_t) \times (R_{t+1}^{nearby} - R_{t+1}^{deferred}) \quad (3.4)$$

$$R_{t+1}^{outright} = \text{sign}(SM_t) \times (R_{t+1}^{nearby}) \quad (3.5)$$

where $R_{t+1}^{spread}(R_{t+1}^{outright})$ is the next month return from the spread (outright) trading strategy and $sign()$ is the sign function.⁸ We compare the returns of the STSM-spread and STSM-outright with the corresponding long-only baseline strategy whose returns are calculated as: $R_{t+1}^{long-spread} = (R_{t+1}^{nearby} - R_{t+1}^{deferred})$ and $R_{t+1}^{long-outright} = (R_{t+1}^{nearby})$, respectively.

3.3. Data and Market Predictability

To facilitate comparison with previous studies, our empirical analysis covers 22 commodities in the energy, metal, grains, and softs categories that are similar to Szymanowska et al. (2014) and Boons and Prado (2019). Table 3.1 lists the commodities and their corresponding exchanges. Daily prices are obtained from the Commodity Research Bureau (CRB) for the period from January 1986 to December 2019. The start of our sample period is the same as Szymanowska et al. (2014), while we extend the sample to end in 2019. The nearest-to-expire contract is defined as the nearby contract. We replace the nearby (first deferred) contract with the first (second) deferred on the first day of the month prior to the contract expiration month.

Table 3.1 presents summary statistics of the spread momentum indicator SM_t (equation 3.3) expressed in annualized percent returns, for each commodity and for the period of analysis. Intuitively, SM_t is a measure of the returns from consistently holding two offsetting positions as a buyer of the nearby and seller of the deferred contract over the past 12 months.⁹ The SM_t average

⁸ In two occasions, $SM_t = 0$. In these cases, the same strategy from the previous month is carried to the next month.

⁹ Since our focus is on spread momentum involving two opposite positions on the same asset, we do not follow the usual practice in outright momentum analyses to subtract the risk-free interest

across markets is -0.01%, with cocoa registering the largest return (4.07%) and the lean hog market the smallest (-3.64%). Negative average returns are found in a majority of markets (15 out of 22), though only 9 are statistically different from zero according to the Welch's t-tests reported in the table. A negative SM_t value signals spread momentum traders to take a short position in the market according to equations (3.4) and (3.5). Positive average returns, recommending a long position in the market, are only identified in 7 markets. However, none of these positive averages is statistically different from zero at the 5% significance level. Volatility, measured as standard error, ranges from 0.03% in gold to 1.59% in lean hog, with the average across all markets being 0.77%. The difference between maximum and minimum returns, another indicator of return dispersion, is minimum in gold (about 6%) and maximum in lean hogs (nearly 61%). A market with a highly volatile SM_t may require rebalancing market positions more often than a market with more stable SM_t .

A necessary condition for spread momentum to exist is predictability: the momentum indicator (SM_t) should be a good predictor of future spread/outright returns in month $t+1$. Here, we assess return predictability before investigating the performance of momentum investment strategies. To assess return predictability, we use two different methods: through regression analysis for each commodity and through a pooled regression including all commodities. The pooled regression provides a joint test of the null hypothesis of no predictability and mitigates the potential data-mining problem from individual market data (Ang and Bekaert, 2007).

3.3.1. *Individual Market Predictability*

For each commodity i we estimate the following regressions by ordinary least squares (OLS),

rate from returns. During the period of our sample, the average risk-free rate (1-year U.S. Treasury Bond) is 2.52%.

$$\begin{bmatrix} R_{t+1}^{long-spread,i} \\ R_{t+1}^{long-outright,i} \end{bmatrix} = \begin{bmatrix} \alpha + \beta SM_t^i + \epsilon_{t+1}^i \\ \gamma + \theta SM_t^i + \eta_{t+1}^i \end{bmatrix}. \quad (3.6)$$

To assess the ability of SM_t^i to predict future returns, we consider the long-only spread (outright) return of market i in month $(t+1)$ as the dependent variable, which we denote as $R_{t+1}^{long-spread,i}$ ($R_{t+1}^{long-outright,i}$). Intuitively, a positive (negative) slope in the regression indicates that the trend in returns over the past 12 months continues (reverts) in the next month $t+1$. Trend reversion casts doubts on the existence of momentum, as the trend is not sustained into the future. The first (second) equation in (3.6) examines whether STSM can predict future spread (outright) returns. We denote the regression as STSM-spread (STSM-outright). Evidence of predictive power supporting spread momentum requires positive and statistically significant slope coefficients (β or θ) and a larger-than-zero R^2 . Results are presented in table 3.2. To compare with prior momentum studies and appreciate the relevance of the sources of bias affecting this literature, we also study predictability including the non-realizable contract roll yield and present results in supplemental table 3.7.

Table 3.2, left panel provides information on the STSM-spread regression slopes, t-statistics, and the goodness-of-fit (adjusted R^2) when the roll yield is excluded;¹⁰ Newey and West (1994) standard errors are used to control for potentially non-*iid* errors.¹¹ We focus on trend continuation

¹⁰ The intercepts are not shown here for the sake of space, but are available on request. A total of 21 of the 22 intercepts are not statistically different from 0. A zero intercept indicates that when current market momentum equals zero, average returns are zero.

¹¹ To derive the Newey and West (1994) standard errors, we use the automatic lag selection technique based on the formula $\text{floor}[4(T/100)^{2/9}]$, where T is the total number of observations. In our case, the lag for each individual market regression is 5 with $T = 408$ (12 months $\times$ 34 years).

which characterizes momentum. Slope coefficients are positive in 8 of the 22 markets in the STSM-spread regression, but only statistically significant in the copper and Minneapolis wheat markets,¹² where the Minneapolis wheat regression slope is 0.011 and the R^2 is 0.127. The Minneapolis wheat futures market is characterized by relatively small volumes. Importantly, during the sample period, the Minneapolis wheat price increased tenfold, and exhibited extreme movement in 2007-8, exhibiting substantial volatility. Previous research findings show that momentum can be driven by both low market liquidity and high volatility (e.g. Asness et al. 2013; Boons and Prado 2019), which may contribute to our regression results. The average slope and R^2 across commodities are -0.001 and 0.058, respectively. The R^2 value suggests the past 12-month returns explain, on average, less than 6% of the variation of the next month return. Also, the negative sign of the slope is suggestive of trend reversal, not continuation.

The right panel of table 3.2 presents coefficients for the STSM-outright regression. We observe 11 positive coefficients, but only 2 are statistically significant — Sugar and West Texas Intermediate oil (WTI), with a slope of 0.002 and 0.004, respectively. Here, possible momentum may be driven by volatility. The average slope coefficient across all markets is virtually zero, with the average R^2 being 0.042. These results are consistent with Tomek (1997) who suggests that outright price forecasting is complex.

Table 3.7 in the supplementary results shows predictability results including non-realizable roll yields. For the STSM-spread (STSM-outright) regression, the average slope coefficient across all markets is 0.38 (0.92), with 11 (10) markets presenting positive and statistically significant slope coefficients. The average R^2 of all individual market regressions is 0.139 (0.051). Average

¹² While the slope in the copper regression is statistically significant, returns are virtually zero.

slope coefficients, t-statistics, and R^2 are larger than the estimations that exclude roll yields, which strongly suggests that previous research on market-level predictability is driven by roll yields.

3.3.2. Pooled Market Predictability

We next study predictability through a pooled regression analysis on all commodities. Pooled regression coefficients can be biased, with the bias coming from three main sources (Ang and Bekaert 2007; Huang et al. 2020). The first bias occurs if identical mean returns are assumed to be the same across different markets. This does not seem appropriate given averages in table 3.1 and requires the use of fixed effects to reduce selection bias. The second source of bias originates from the momentum literature's common practice to scale returns using their past volatility. Since volatility differs dramatically across assets, volatility scaling without fixed effects can exacerbate the bias. Third, the use of the persistent past 12-month return (SM_t^i) as a predictor can cause relevant size distortions. Combined, the three sources usually lead to over-rejection of the null hypothesis that the slope is not statistically different from zero (Ang and Bekaert 2007; Huang et al. 2020).

We run the pooled regression both under equal and volatility weighting. In the latter case, we normalize returns using each commodity's *ex-ante* volatility. We follow Moskowitz et al. (2012) and Huang et al. (2020) and calculate the variance σ_{t-12}^2 as the exponentially weighted squared daily return variance in the past 12 months (121 days) as:

$$\left[\begin{array}{c} (\sigma_{t-12}^{long-spread,i})^2 \\ (\sigma_{t-12}^{long-outright,i})^2 \end{array} \right] = \left[\begin{array}{c} 261 \sum_{j=0}^{121} (1-\delta)\delta^j \left(R_{t-1,-j}^{long-spread,i} - \bar{R}_{t-12}^{long-spread,i} \right)^2 \\ 261 \sum_{j=0}^{121} (1-\delta)\delta^j \left(R_{t-1,-j}^{long-outright,i} - \bar{R}_{t-12}^{long-outright,i} \right)^2 \end{array} \right] \quad (3.7)$$

where j represents days; for example, $R_{t-1,-j}^{long-outright,i}$ represents the return of the last day of month $t-2$ (month $t-1$ minus 1 day). The scalar 261 in equation (3.7) annualizes the variance. Parameter δ is chosen such that the span is 121 days, then $\sum_{j=0}^{121} (1-\delta)\delta^j \times j \cong \frac{\delta}{1-\delta} = 60$. Expression

$\bar{R}_{t-12}^{long-spread,i}$ ($\bar{R}_{t-12}^{long-outright,i}$) is market i 's exponentially weighted average spread (outright) return in the past 12 months (121 days) computed as:

$$\begin{bmatrix} \bar{R}_{t-12}^{long-spread,i} \\ \bar{R}_{t-12}^{long-outright,i} \end{bmatrix} = \begin{bmatrix} 261 \sum_{j=0}^{121} (1-\delta) \delta^j R_{t-1,-j}^{long-spread,i} \\ 261 \sum_{j=0}^{121} (1-\delta) \delta^j R_{t-1,-j}^{long-outright,i} \end{bmatrix} \quad (3.8)$$

The pooled regressions with volatility scaling become the following:

$$\begin{bmatrix} R_{t+1}^{long-spread,i} / \sigma_{(t+1)-12}^{long-spread,i} \\ R_{t+1}^{long-outright,i} / \sigma_{(t+1)-12}^{long-outright,i} \end{bmatrix} = \begin{bmatrix} \alpha + \beta (SM_t^i / \sigma_{t-12}^{long-spread,i}) + \epsilon_{t+1}^i \\ \gamma + \theta (SM_t^i / \sigma_{t-12}^{long-spread,i}) + \eta_{t+1}^i \end{bmatrix} \quad (3.9)$$

To correct for the differences in asset mean returns, we demean spread and outright returns from each market in equation (3.10), which is equivalent to using fixed effects:

$$\begin{bmatrix} R_{t+1}^{long-spread,i} / \sigma_{(t+1)-12}^{long-spread,i} - \overline{R_{(t+1)}^{long-spread,i} / \sigma_{(t+1)-12}^{long-spread,i}} \\ R_{t+1}^{long-outright,i} / \sigma_{(t+1)-12}^{long-outright,i} - \overline{R_{(t+1)}^{long-outright,i} / \sigma_{(t+1)-12}^{long-outright,i}} \end{bmatrix} = \begin{bmatrix} \alpha + \beta \left(SM_t^i / \sigma_{t-12}^{long-spread,i} - \overline{SM_t^i / \sigma_{t-12}^{long-spread,i}} \right) + \epsilon_{t+1}^i \\ \gamma + \theta \left(SM_t^i / \sigma_{t-12}^{long-spread,i} - \overline{SM_t^i / \sigma_{t-12}^{long-spread,i}} \right) + \eta_{t+1}^i \end{bmatrix}. \quad (3.10)$$

Table 3.3 reports the pooled regression coefficients before and after controlling for the differences in means and volatilities.¹³ We also report regression coefficients before and after including the roll yield. The STSM-spread slope coefficient is 0.139 and statistically significant at 1% level if we exclude the roll yield, but do not consider fixed effects and volatility scaling (regression (t3.1)). Hence, on average, a 1% increase in past 12-month spread returns leads to a 0.139% increase in the next month spread returns, suggesting potentially profitable trend following trading strategies. Our slope coefficient from equation (t3.1) is much smaller than the equivalent

¹³ We focus on the slope coefficient of the regressions. None of the intercepts are statistically significant in the 12 regressions presented here.

slope of 2.34 reported in Boons and Prado (2019). Regression (t3.2) shows that after controlling for volatility differences across markets, the coefficient becomes negative, signaling trend reversion and both statistical significance of the slope and R^2 decline. When controlling for both differences in volatilities and means (regression (t3.3)), evidence of predictability disappears completely with R^2 dropping to 0. When including the non-realizable roll yield, STSM-spread regression results provide much stronger evidence of spread momentum with larger t-statistics and R^2 (regressions (t3.4) to (t3.6)). The magnitude of the slope coefficient in equation (t3.4), i.e., 2.018, is close to what Boons and Prado (2019) find. Our R^2 coefficient of 2.6% is also close to the 2% in Boons and Prado (2019). After volatility scaling and inclusion of fixed effects, slope coefficients and R^2 are still relatively large compared to the previous equations without roll yield ((t3.1) to (t3.3)).

The bottom half of table 3.3 presents pooled regression results for the STSM-outright regression. Without the non-realizable roll yield, we find the slope coefficient to be 0.451 (t3.7) before controlling for return volatility and mean differences across markets. This suggests that a 1% increase in past 12-month spread returns leads to a 0.451% increase in the next month outright returns. The zero slope coefficients in (t3.8) and (t3.9) suggest that the slope in (t3.7) is a reflection of econometric misspecifications. When including the non-realizable roll yield, we find a 1% increase in past 12-month spread return is associated with a 9.372% increase in the outright return (t3.10), with an R^2 of 1.2%. Our findings from (t3.10) are similar to what Boons and Prado (2019) found (slope = 10.45, R^2 = 1%). After volatility scaling and inclusion of fixed effects, we find much smaller and statistically insignificant slope coefficients ((t3.11) and (t3.12)). The R^2 values also drop to zero.

When including the roll yield and without using volatility scaling and fixed effects, our results in table 3.3 ((t3.4) and (t3.10)) are similar to the Boons and Prado (2019) results. However, the estimated slope coefficients are much smaller without the roll yields for both STSM-spread and STSM-outright models. Once again, we find evidence suggesting that previous findings are substantially driven by inclusion of the non-realizable roll yields. Moving forward, we concentrate primarily on the realizable returns, excluding roll yields. These provide a more realistic measure of attainable market opportunities. While table 3.3 strengthens the evidence against return predictability, it does not use robust critical values to test for predictability. We now assess robustness of predictability with robust critical values.

3.3.3. *Bootstrap Tests*

Due to concerns of statistical bias leading to over-rejection of the null hypothesis of no predictability ($\beta = 0$, $\theta = 0$), we use nonparametric bootstrapping techniques. Our bootstrap approach resamples with replacement from the real data to draw $T = 8,976$ pairs¹⁴ of the STSM-spread (STSM-outright) regression dependent and independent variables. After each draw, the pooled regression is re-estimated and a new $\hat{\beta}$ ($\hat{\theta}$) and corresponding t-statistic are produced from equation (3.10). This pair bootstrap procedure does not preserve any serial correlation in the data (Kreiss and Lahiri 2012; Zakamulin and Giner 2020). The bootstrap allows us to identify proper critical values for the t-test that accommodate heteroscedasticity and data-generating processes that are more general than those based on the parametric approaches. The procedure is repeated 1,000 times, generating an array of 1,000 new t-statistics. We define the 5% significance level of the t-statistic as the 97.5% quantile from the set of simulated t-statistics. If the original pooled

¹⁴ T is the total number of observations in the pooled data (22 markets x 34 years x 12 months/year).

regression t-statistic is larger than the usual 5% t-critical value (1.96) but smaller than the simulated critical value, we conclude that the pooled regression tends to over-reject the null and the statistical evidence of predictability is weak.

We apply the pair bootstrap method to regression specifications (t3.3) and (t3.9) in table 3.3 and present results in table 3.4 top and bottom panels, respectively. Both regressions control for volatility and fixed effects and exclude the non-realizable roll yield. As a robustness check, we investigate predictability with different look-back months ($L=1$ to 12). The top panel shows that the STSM-spread regression (t3.3) only produces four positive coefficients of the 12 look-back months, though only one ($L=1$) is positive and statistically significant under the t-distribution. However, the slope coefficient for $L=1$ is not statistically significant under the bootstrapped distribution, suggesting that over-rejection occurs. Eight coefficients are negative, with absolute values very close to and not statistically different from zero under the bootstrap distribution. In one scenario, we find the slope coefficient to be negative and statistically significant under the t-distribution ($L = 7$), which suggests potential trend reversion instead of continuation.

Consistent with the top panel, the bottom panel in table 3.4 shows that spread momentum does not lead to outright return predictability. The t-distribution leads to over-rejection of the null hypothesis of no return predictability at $L=1$ in the STSM-outright regression. Slopes and R^2 coefficients are generally lower in the bottom panel, suggesting that spread momentum is a better predictor of spread than outright returns. Our empirical evidence here contrasts with previous studies (e.g. Szymanowska et al. 2014, and Boons and Prado 2019) who conclude that spread market momentum is an effective predictor of future outright returns. Consistent with Huang et al. (2020), statistical evidence to support spread momentum is weak across different look-back months.

3.4. Momentum Trading Profitability

These results provide, at best, only weak statistical evidence of commodity return predictability, which does not support the existence of spread momentum. We now turn to the profitability of momentum trading in both individual markets and a portfolio of commodities.

We compare the profitability of STSM with a long-only benchmark strategy. As opposed to momentum investment, the long-only strategy does not require return predictability. A comparison between momentum and long-only strategies is of particular interest to financial funds that have been shown to use both strategies (Kang et al. 2020). Equation (3.11) demonstrates the difference between the STSM and long-only strategies in spreads and outright futures:

$$\begin{bmatrix} R_{t+1}^{STSM-spread,i} \\ R_{t+1}^{STSM-outright,i} \\ R_{t+1}^{Long-spread,i} \\ R_{t+1}^{Long-outright,i} \end{bmatrix} = \begin{bmatrix} sign(SM_t) \times (R_{t+1}^{nearby,i} - R_{t+1}^{deferred,i}) \\ sign(SM_t) \times (R_{t+1}^{nearby,i}) \\ (R_{t+1}^{nearby,i} - R_{t+1}^{deferred,i}) \\ (R_{t+1}^{nearby,i}) \end{bmatrix}. \quad (3.11)$$

We use the nonparametric Wilcoxon test to compare the differences in returns between the STSM and benchmark strategies.

Table 3.5 presents annualized returns from the STSM and long-only strategies for each market. During our sample period, the STSM-spread strategy return averaged 0.28%, which is higher than but not statistically different from the static long-only spread strategy return (-0.01%). The return is also substantially smaller than the average 2.52% (1-year U.S. Treasury Bond) risk-free return over the sample period. The average return from the STSM-outright strategy is -1.86%, similar to the simple long-only outright return (-1.83%). These results are consistent with table 3.4, showing that STSM momentum is a better predictor of spread than outright returns. Among the 4 strategies from equation (3.11), the STSM-outright strategy shows the largest average losses, suggesting

some risks to momentum trading. We examine the distribution of momentum strategy returns below. Our STSM-spread and STSM-outright strategy returns are smaller than the return in outright markets shown in Huang et al. (2020), which may well reflect the fact that these authors use non-realizable roll yields.

For individual markets, STSM-spread returns range from 6.70% in the sugar market to -6.86% in the lean hog market, with 10 markets presenting positive average returns. While STSM-spread seems to outperform the long-only spread in 10 markets, return differences are not statistically significant with the exception of gold. While the STSM-spread strategy outperforms the long-spread by 0.63% in gold, both strategies generate negative returns in this market. In contrast, the long-only spread strategy returns are statistically larger than the STSM-spread returns in 8 of the 22 markets with 5 of them generating positive long-only returns. For the STSM-outright strategy, coffee has the lowest return (-9.35%) and soybean has the largest (6.63%). The long-only outright returns are statistically significant under the 5% significance level and larger than the STSM-outright returns in only one market: coffee. While STSM-outright returns are larger than the long-only return in ten markets, none of the differences is statistically significant. Overall, our results indicate that STSM returns vary significantly across individual markets and the STSM strategy tends to underperform the simple long-only strategy.

Table 3.8 in the supplementary results presents individual market STSM returns including the contract roll yield. The average STSM-spread (STSM-outright) return of 5.56% (5.45%) is much larger than realizable 0.28% (-1.86%) return shown in table 3.5. This provides further evidence that the unrealizable roll yield plays an important role in previous STSM profitability analyses.

Prior research has suggested that positive returns may be generated through portfolio diversification, despite the fact that each individual market has an expected return of zero, which

is referred to as the “diversification return” (Gorton and Rouwenhorst 2006; Willenbrock 2011; Boons and Prado 2019). As a result, we create a portfolio including all commodities and apply the STSM and long-only investment strategies to examine whether a diversified portfolio increases total returns. The portfolios are built under both equal-weighting and volatility-weighting schemes. Table 3.6 shows STSM portfolio arithmetic average returns and their difference from long-only returns with different look-back month strategies.¹⁵ With very few exceptions, STSM and long-only investment strategies fail to generate statistically different returns. On average, the STSM-spread equal-weighting (volatility-weighting) portfolios generate an annualized return on the order of -0.49% (-1.64%) outperforming (underperforming) the spread long-only strategy by 0.37% (0.78%). Returns range from -2.19% to 0.39% (-4.50% to 0.72%) for the STSM-spread equal-weighting (volatility-weighting) portfolio. Notice that the maximum gains are not compensated by the maximum losses, suggesting potential downside risks. Similar results emerge in the outright markets. On average, we find the STSM-outright strategy underperforms the outright long-only strategy by 3.12% (2.48%) under equal-weighting (volatility-weighting). Given larger volatility experienced by outright markets relative to spreads, our results suggest that momentum strategies suffer much higher downside risks in outright markets than spreads. The STSM-outright returns fluctuate from -12.64% to 3.11% (-6.83% to 2.66%) with equal-weighting (volatility-weighting), which again indicates large downside risks. In some cases, we find the STSM-outright strategy to

¹⁵ Sanders and Irwin (2012) point out that portfolio diversification return can stem from using the geometric average returns, and the diversification return is an increasing function of asset return variance and number of assets. To avoid this estimation bias and similar to earlier momentum studies (e.g. Moskowitz et al. 2012; Huang et al. 2020), we calculate the arithmetic average returns.

underperform the long-only outright strategy substantially. For example, the average STSM-outright return at $L = 3$ is 14.47% lower than the long-only return under equal-weighting. With volatility-weighting, the STSM-outright still underperforms as much as 8.66% ($L = 2$).

The STSM returns here are smaller than previous literature findings (Miffre and Rallis 2007; Asness et al. 2013; Georgopoulou and Wang 2017), including those in Boons and Prado (2019) and Szymanowska et al. (2014) who adopt smaller portfolios and more stringent filtering techniques. We conduct additional analyses to shed light on the role of roll yield in portfolio returns (table 3.9 in the supplementary results) and returns become positive and substantially larger than the ones reported here. Consistent with the predictability analysis and commodity-level STSM investment strategies, this suggests that previous research overestimated returns by including the roll yields. Overall, the results in tables 3.5 and 3.6 provide no support for the proposition that momentum trading strategies can generate significant positive returns in the commodity futures markets.

3.4.1. When Would Momentum Work and How Risky Would It Be?

As discussed earlier, our results suggest that STSM strategies do not outperform long-only strategies (most differences in returns are not statistically significant), which is consistent with evidence provided earlier of no return predictability. Here, we conduct a simulation analysis that identifies first, the degree of return predictability required for the STSM-spread/outright to outperform the long-only spread/outright strategy and second, the risks of the STSM profitability.

Following Huang et al. (2020), we assume a data-generating process based on equation (3.10). To facilitate interpretation of the slope coefficient, however, the exogenous variable is divided by 12 so that it represents the monthly average return in the past 12 months. The regression is estimated using the pooled monthly portfolio returns from the real data which yields $\beta = 0.0091$

and $\theta = 0.0016$. We impose different degrees of predictability by setting $\beta(\theta)$ to different values, $\beta(\theta) = 0.0182$ (0.0032) and 0.0273 (0.0048). These are twice and three times as large as the $\beta(\theta)$ estimated from real data. Economically speaking, $\beta = 0.0273$ ($\theta = 0.0048$) indicates that a 1 percentage point increase in the monthly average spread return in the past 12-months is associated with a 0.0273% (0.0048%) increase in the next 1-month spread (outright) return. Based on the estimated parameters, we apply a wild bootstrap technique to simulate the residual series and the dependent variable. We then take the bootstrapped samples and compare the profitability of STSM and long-only strategies.

Since the objective of this bootstrap is to simulate the data based on a fixed level of predictability, the nonparametric bootstrap is not applicable. Hence, we apply a parametric wild bootstrap method that introduces *i.i.d.* shocks into the residuals and mimics the dependence properties of the observations (Kreiss and Lahiri 2012). The wild bootstrap first generates T elements (v_{t+1}^i) randomly drawn from a 2-point Rademacher distribution, with mean 0 and variance 1.

$$\begin{bmatrix} v_{t+1}^{spread,i} \\ v_{t+1}^{outright,i} \end{bmatrix} = \begin{cases} 1, & \text{with probability of } 1/2 \\ -1, & \text{with probability of } 1/2 \end{cases} \quad (3.12)$$

Then, a new residual series is produced as $v_{t+1}^{spread,i} \times \hat{\epsilon}_{t+1}^i$ ($v_{t+1}^{outright,i} \times \hat{\eta}_{t+1}^i$), where $\hat{\epsilon}_{t+1}^i$ and $\hat{\eta}_{t+1}^i$ are the fitted residuals from equation (3.10). A new set of dependent variables is generated from:

$$\begin{bmatrix} \left(R_{t+1}^{long-spread,i} / \sigma_{(t+1)-12}^{long-spread,i} - \overline{R_{(t+1)}^{long-spread,i} / \sigma_{(t+1)-12}^{long-spread,i}} \right)^* \\ \left(R_{t+1}^{long-outright,i} / \sigma_{(t+1)-12}^{long-outright,i} - \overline{R_{(t+1)}^{long-outright,i} / \sigma_{(t+1)-12}^{long-outright,i}} \right)^* \end{bmatrix} = \begin{bmatrix} \hat{\alpha} + \hat{\beta} \left(SM_t^i / \sigma_{t-12}^{long-spread,i} - \overline{SM_t^i / \sigma_{t-12}^{long-spread,i}} \right) / 12 + v_{t+1}^{spread,i} \hat{\epsilon}_{t+1}^i \\ \hat{\gamma} + \hat{\theta} \left(SM_t^i / \sigma_{t-12}^{long-spread,i} - \overline{SM_t^i / \sigma_{t-12}^{long-spread,i}} \right) / 12 + v_{t+1}^{outright,i} \hat{\eta}_{t+1}^i \end{bmatrix}, \quad (3.13)$$

where * indicates the value is simulated from the bootstrap. The wild bootstrap procedure is repeated 1,000 times. For each pseudo-sample we calculate the returns from implementing STSM-spread (STSM-outright) investment strategies. This generates 8,976 monthly returns (22 markets $\times$ 34 years $\times$ 12 months) for every pseudo-sample. We then compare returns against their respective benchmark strategies.

Plots in figure 3.1 show the average return distributions of the annualized returns obtained under STSM-spread, given different beta values. For $\beta = 0.0182$, top panel, although distributions overlap, the long-spread strategy performs slightly worse than the STSM-spread strategy (mean returns are 0.503% and 0.76%, respectively). For $\beta = 0.0273$, bottom panel, the STSM-spread strategy outperforms the long-spread strategy (mean returns are centered at 1.65% and 0.60%, respectively). Despite differences in mean returns for the two strategies, the risk that each strategy implies is still unclear. To shed light on risk, we focus on $\beta = 0.0273$ and for each of the 1,000 simulated samples, we take the 90% and 10% quantile of the monthly annualized returns. We plot their distributions in figure 3.2 top and bottom panels, respectively. At the 90% quantile, the STSM-spread (long-spread) profit is centered at 4.41% (2.79%), with the former being 36% larger than the latter. The 90% quantile distribution indicates that the STSM-spread strategy has more potential to generate large returns than the long-spread. However, the 10% STSM-spread (long-spread) quantile return is centered at -13.08% (-7.14%), with the loss under the STSM strategy being 45% larger than the loss under the long-only strategy.

Figure 3.3 shows the monthly average of the annualized return distributions for the STSM-outright and long-outright strategies. Similarly, we generate the distributions with $\theta = 0.0032$ (0.0048), which is twice (three times) stronger than the estimated coefficient from real data and present results in the top (bottom) panel. Despite return distributions overlap when $\theta = 0.0032$, the

STSM-outright generates less losses than the long-outright, with average returns of -0.34% and -0.62%, respectively. For $\theta = 0.0048$, the STSM-outright strategy is clearly superior to the long-outright alternative, with average returns of 3.65% and 0.94%, respectively. The top panel in figure 3.4 demonstrates the distribution for 90% quantile returns under $\theta = 0.0048$. Similar to the STSM-spread strategy, we find the STSM-outright strategy offers larger upside returns compared to long-only strategies. However, the downside risk in STSM-outright is also substantial. The bottom panel of figure 3.4 shows the 10% quantile return for STSM-outright (long-outright) is centered at -23.08% (-14.36%), which implies that the STSM-outright strategy faces 37.7% higher downside losses than the static long-outright strategy in commodity futures markets. These findings are consistent with Daniel and Moskowitz (2016) who suggest that mean returns generated from momentum strategies tend to be negatively skewed. This implies that momentum strategies are exposed to larger downside risks and negative returns can be pronounced.

3.5. Conclusion

Some published research has concluded that past returns can predict future returns in a wide array of financial markets. This predictive ability indicates the possibility of producing significant returns from a time series momentum (TSM) trading strategy, which challenges the market efficiency hypothesis (Fama 1970). TSM returns have been found to be larger than market risk-free returns, even after controlling for different risk levels across different markets to ensure that returns are not dominated by high-volatility assets (Moskowitz et al. 2012; Asness et al. 2013; Szymanowska et al. 2014; Boons and Prado 2019). Such findings have also been claimed for futures markets, where evidence of both outright and spread time series momentum would point to persistent mispricing. If accurate, these findings should have important implications for various groups of market participants. For investors, a strong market momentum implies profit

opportunities that have not been arbitrated away and can be realized following a simple trading rule. Persistent trending and mispricing however, cast doubt on the ability of the futures market to discover prices. This has implications for those who rely on futures markets to hedge and inform their decisions. Biased futures' price forecasts bring economic inefficiencies, may reduce the presence of commercials in futures markets and eventually threaten existence of the market (Working 1970, Leuthold et al. 1994). Evidence of market inefficiency is also relevant to the regulators that promote the integrity, resiliency and vibrancy of the markets (CFTC 2020).

Recently, several studies have reexamined the existence of TSM, correcting for econometric and empirical design issues that were pervasive in earlier work and may have led to misleading conclusions. While informative, these recent studies have not addressed all of the design and statistical issues with TSM. Further, no research has yet questioned existing evidence on spread time series momentum (STSM). In this article, we examine whether STSM leads to high profitability in both spread (STSM-spread) and outright (STSM-outright) investment strategies. We use a sample of 22 U.S. commodity futures markets observed from 1986 to 2019. For the first time in the momentum literature, we adopt a comprehensive approach that corrects for both the empirical design and econometric issues affecting previous research.

To ensure that reported investment returns are realizable (Bessembinder 2018; Irwin et al. 2020; Paschke et al. 2020), we exclude the futures contract indicative but unobtainable roll yield when calculating the momentum strategies' profits or losses. Since STSM profitability is predicated on return predictability, we use both individual market and pooled regression analysis to shed light on this matter. We regress future spread/outright returns using spread returns during the past several months as a predictor. We do not follow the usual practice of imposing a common intercept in the pooled regression which implies that mean returns are identical across different

markets. This stringent assumption may cause upward biased regression slope coefficients, a bias that could be exacerbated under the common practice of scaling returns by volatility and the use of a persistent predictor such as ours. Upward bias will result in over-rejection of the null hypothesis of no predictability. We correct for these statistical issues and conduct bootstrap simulations to produce more accurate estimates and tests for the null hypothesis of no predictability. Empirical results demonstrate that, although evidence of return predictability can be found, it is mainly driven by inclusion of unrealizable roll yields, with statistical issues being less relevant.

We then study returns from STSM-spread and STSM-outright investment strategies. On average, returns are very small and mostly negative, suggesting that no persistent trending exists in futures markets over the sample period. We use the static long-only strategy as a benchmark that does not require predictability. We compare STSM returns against this benchmark and find that, generally, there are no statistically significant differences between the two strategies, with both alternatives generating very low or negative returns. The few significant differences suggest the benchmark performs better than the STSM strategies. Based on simulation analyses, we show that for the STSM strategy to outperform the long-only benchmark strategy, the spread predictability needs to be three times stronger than observed from the real data. Although STSM strategies could potentially generate higher mean returns under stronger market predictability, the downside risks of STSM-spread (STSM-outright) trading can be as much as 45% (38%) larger than the long-only spread (outright) benchmark strategy. As momentum trading strategies become more popular among managed funds, our results show they are unlikely to be more profitable than long-only strategies and may entail substantially larger downside risks. Low profitability of STSM-investment strategies suggests that futures markets have operated efficiently throughout

our sample period. Hence, commercials and regulators should not worry about market inefficiencies related to time series momentum.

Our regression results are consistent with results in Huang et al.'s (2020) for the outright-futures markets, but contrast with several prominent momentum studies (e.g. Moskowitz et al. 2012; Asness et al. 2013; Szymanowska et al. 2014; Boons and Prado 2019). Our analysis strongly suggests that the differences largely reflect the inclusion of the unrealizable roll yield, with statistical flaws playing a lesser role. Finally, while some research suggests that returns to a portfolio may be positive even when individual market returns are zero or negative (e.g. Gorton and Rouwenhorst 2006; Willenbrock 2011; Boons and Prado 2019), our empirical findings in the context of STSM do not support the “diversification return” in commodity futures markets.

Several topics are of interest for future research. First, we do not consider the effect of transaction costs while implementing the momentum strategies. This could potentially eliminate any remaining evidence of momentum profits. Second, a thorough examination of time-varying predictability may bring a deeper understanding of price behavior and momentum trading, especially in the context of fast electronic markets, where price inefficiencies may be arbitrated away much faster than before. It is still unclear whether positive momentum profit is a reflection of time-varying risk premium for bearing larger risks, which offers scope for future inquiry.

3.6. Tables and Figures

Table 3.1: Summary Statistics of the Spread Momentum Indicator, SM_t (1986-2019)

	Exchange	Mean	P-Value	Volatility (Std. Error)	Min.	Max.
Corn	CBOT/CME	-1.94%	0.00	0.67%	-10.31%	8.18%
Soybeans	CBOT/CME	-1.54%	0.53	0.41%	-8.03%	6.73%
Wheat-KC	KCBOT/CME	2.94%	0.13	0.64%	-10.01%	9.43%
Wheat-Chi	CBOT/CME	-1.84%	0.00	0.69%	-11.87%	11.25%
Wheat-MN	MGEX	-0.25%	0.05	0.86%	-11.23%	13.88%
Rough Rice	CBOT/CME	-1.93%	0.00	0.96%	-10.87%	8.74%
Soy-meal	CBOT/CME	-0.24%	0.03	0.68%	-9.22%	12.79%
Soy-oil	CBOT/CME	2.46%	0.09	0.27%	-6.39%	4.13%
Oats	CBOT/CME	1.67%	0.33	1.38%	-19.57%	21.33%
Live Cattle	CBOT/CME	-1.06%	0.04	0.80%	-14.49%	17.16%
Lean Hog	CBOT/CME	-3.64%	0.00	1.59%	-28.51%	32.43%
Feed. Cattle	CBOT/CME	3.61%	0.21	0.49%	-10.14%	9.63%
Cocoa	NYBOT/ICE	4.07%	0.15	0.47%	-9.58%	8.01%
Sugar	NYBOT/ICE	-1.09%	0.76	1.32%	-21.60%	24.46%
Coffee	NYBOT/ICE	2.87%	0.88	1.17%	-9.11%	5.61%
Cotton	NYBOT/ICE	-2.46%	0.01	0.97%	-15.54%	15.87%
Lumber	CBOT/CME	-2.33%	0.00	1.46%	-23.54%	28.53%
Orange Juice	NYBOT/ICE	-0.77%	0.42	0.73%	-11.86%	14.33%
WTI	NYMEX/CME	-1.28%	0.57	0.90%	-11.98%	11.37%
Gold	COMEX/CME	-0.66%	0.25	0.03%	-4.03%	2.12%
Copper	COMEX/CME	3.90%	0.21	0.26%	-5.25%	3.51%
Silver	COMEX/CME	-0.61%	0.18	0.20%	-4.88%	2.46%
Average		-0.01%	0.21	0.77%	-12.18%	12.36%

Notes: Columns 1 and 2 identify the commodities and the exchanges on which each commodity is traded. Column 3 provides the arithmetic mean of the spread momentum indicator SM_t (equation (3.3) in the article), expressed in annualized returns. Column 4 is the p-value of the Welch's t-test to examine whether the mean SM_t is statistically different from zero. Column 5 reports the standard error of SM_t . Columns 6 and 7 report the minimum and maximum monthly SM_t , respectively.

Table 3.2: Spread Momentum Predictability (1986-2019, 12-Month Look-back Period)

	<i>STSM-spread</i>			<i>STSM-outright</i>		
	Slope	t-stat. (slope)	R ²	Slope	t-stat. (slope)	R ²
Corn	-0.002	0.037	0.007	-0.003	0.166	0.005
Soybeans	0.000	0.129	0.063	0.003	1.052	-0.002
Wheat-KC	-0.001	1.107	0.041	-0.002	0.446	0.027
Wheat-Chi	-0.003	0.061	0.013	-0.004	0.429	0.030
Wheat-MN	0.011**	2.137	0.127	0.001	0.330	0.049
Rough Rice	-0.003	0.049	0.010	-0.005	0.110	0.002
Soy-meal	0.001	0.180	0.081	0.006	0.397	0.016
Soy-oil	-0.001	0.108	0.032	-0.002	0.930	0.020
Oats	-0.002	1.120	0.051	-0.003	0.506	0.102
Live Cattle	-0.001	0.162	0.069	0.002	0.235	0.034
Lean Hog	-0.007	0.185	0.064	-0.004	0.429	0.095
Feed. Cattle	-0.001	0.261	0.101	0.003	0.633	0.045
Cocoa	-0.001**	2.206	0.088	-0.003	1.167	0.051
Sugar	-0.001	0.084	0.019	0.002**	2.370	0.046
Coffee	0.013	0.238	0.175	-0.007	0.927	0.152
Cotton	-0.002	1.028	0.007	0.000	0.040	-0.001
Lumber	-0.004	0.224	0.108	-0.006	0.628	0.143
Orange Juice	0.000	0.151	0.066	-0.002	0.451	0.021
WTI	-0.001	0.170	0.084	0.004*	1.681	0.069
Gold	0.000	0.055	0.024	0.002	0.604	-0.001
Copper	0.000***	3.123	0.032	0.006	1.264	0.020
Silver	0.000	0.022	-0.002	0.009	0.470	-0.006
Average	-0.001	0.583	0.058	0.000	0.693	0.042

Notes: This table reports results of the market-level regression analysis that investigates the ability of the spread momentum indicator (SM_t^i) to predict future spread (outright) returns (equation (3.6) in the paper). The dependent variable is the spread (outright) return, $R_{t+1}^{long-spread,i}$ ($R_{t+1}^{long-outright,i}$) for month $t+1$. The independent variable is SM_t^i calculated over the past 12-month as in equation (3.3) in the paper. The non-realizable roll yields are excluded from the returns' calculation. There are $T=408$ (12 months $\times$ 34 years) observations for each market. Column 2 (5) shows the regression slope for spread (outright) return predictability. The corresponding t-statistics in column 3 (6) are based on Newey and West (1994) standard errors to control for non-*iid* errors. In applying Newey-West, we use the automatic lag selection technique formula ($\text{floor}[4(T/100)^{2/9}]$). In our case, the lag for each individual market regression is 5. ***, **, * indicate significance at the 1%, 5% and 10% levels.

Table 3.3: Pooled Regression Spread Momentum Predictability (1986-2019, 12-Month Look-back Period)

Regression	Volat. Weight	Fixed Effects	Slope	t-stat. (slope)	R ²
Without Roll Yield (STSM-spread)					
(t3.1)	No	No	0.139***	7.939	0.014
(t3.2)	Yes	No	-0.001	-0.196	0.008
(t3.3)	Yes	Yes	0.000	0.246	0.000
With Roll Yield (STSM-spread)					
(t3.4)	No	No	2.018***	4.11	0.026
(t3.5)	Yes	No	0.711*	1.803	0.017
(t3.6)	Yes	Yes	0.208**	2.121	0.055
Without Roll Yield (STSM-outright)					
(t3.7)	No	No	0.451***	7.757	0.008
(t3.8)	Yes	No	-0.000	-0.057	-0.000
(t3.9)	Yes	Yes	0.000	0.039	-0.000
With Roll Yield (STSM-outright)					
(t3.10)	No	No	9.372***	6.410	0.012
(t3.11)	Yes	No	0.072	0.062	-0.000
(t3.12)	Yes	Yes	0.010	0.055	-0.000
Notes: This table reports results of the pooled regression analysis that investigates the ability of the spread momentum indicator (SM_t^i) to predict future spread (outright) returns (equation (3.6), (3.9), and (3.10) in the paper). The dependent variable is the spread (outright) return, $R_{t+1}^{long-spread,i}$ ($R_{t+1}^{long-outright,i}$) for month $t+1$. The independent variable is SM_t^i , calculated over the past 12-months as in equation (3.3) in the paper. There are T=8976 (12 months $\times$ 34 years $\times$ 22 markets) observations for each regression. The first two numerical columns show the slope in each regression, and corresponding t-statistics based on Newey and West (1994) standard errors to control for non- <i>iid</i> errors. In applying Newey-West, we use the automatic lag selection technique formula (floor[4(T/100) ^{2/9}]). In our case, the lag for the pooled regression is 10. The adjusted R ² is displayed in the last column. ***, **, * indicate significance at the 1%, 5% and 10% levels.					

Table 3.4: Volatility and Fixed Effects Adjusted Pooled Regression with Different Look-back Periods and Bootstrapped t-Statistics (1986-2019)

	Non-bootstrap Regression			Bootstrap
Look-back Month (L)	Slope	t-statistic (slope)	R ²	Critical Value
STSM-spread (regression t3.3)				
1	0.924	4.990	0.072	8.791
2	-0.004	-0.458	0.035	0.542
3	0.002	0.072	0.000	5.210
4	-0.002	-1.094	0.018	0.794
5	0.001	0.043	0.001	13.461
6	-0.001	0.042	0.001	17.178
7	-0.002	-0.576	0.025	1.182
8	-0.003	-1.016	0.015	1.638
9	-0.003	-1.377	0.025	0.608
10	-0.003	-7.444	0.016	1.402
11	-0.003	-0.526	0.013	2.079
12	0.000	0.246	0.000	0.147
STSM-outright (regression t3.9)				
1	1.709×10^{-3}	2.553	0.002	4.942
2	-4.298×10^{-6}	-0.127	0.000	0.832
3	4.319×10^{-5}	0.340	0.000	2.130
4	-5.143×10^{-7}	-0.043	0.000	1.038
5	1.976×10^{-5}	0.251	0.000	8.773
6	2.020×10^{-5}	0.213	0.000	9.535
7	-1.423×10^{-6}	-0.093	0.000	0.922
8	1.145×10^{-6}	0.053	0.000	1.510
9	-1.835×10^{-6}	-0.181	0.000	0.847
10	1.030×10^{-6}	0.049	0.000	1.602
11	-1.616×10^{-6}	0.055	0.000	1.688
12	1.319×10^{-4}	0.039	0.000	0.942

Table 3.4 (cont.)

Notes: This table shows the pooled regression coefficients of the spread momentum predictability for regressions (t3.3) and (t3.9) in table 3.3. Both regressions control for volatility and fixed effects (equation (3.10) in the paper). The dependent variable is the 1-month spread (outright) return $R_{t+1}^{long-spread,i}$ ($R_{t+1}^{long-outright,i}$) for STSM-spread (STSM-outright) from month $t+1$. The independent variable is SM_t^i calculated as in equation (3.3) in the article with different look-back months (L, first column). A constant is included in the regression. The second and third columns show the slope coefficient and the corresponding t-statistics in each regression based on Newey and West (1994) standard errors to control for non-*iid* errors. Due to the magnitude of the coefficients, we use the scientific notation to present the slope coefficients for regression (t3.9). The goodness of fit (adjusted R^2) is displayed in the fourth column. The last column shows the values of the 97.5% quantile t-statistics generated from the 1,000 pair bootstrap simulations.

Table 3.5: STSM and Long-Only Trading Strategies Annualized Return Comparisons (Individual Markets)

	STSM-spread	STSM-Spread vs. Long-Spread			STSM-outright	STSM- Outright vs. Long-Outright		
		Long Spread	Diff.	p-value		Long Outright	Diff.	p-value
Corn	-4.70%	-1.94%	-2.76%	0.00	-2.48%	-5.11%	2.63%	0.24
Soybeans	-0.33%	-1.54%	1.21%	0.48	6.63%	6.56%	0.07%	0.23
Wheat-KC	2.36%	2.94%	-0.58%	0.30	3.14%	-4.01%	7.15%	0.39
Wheat-Chi	1.41%	-1.84%	3.25%	0.36	-9.15%	-5.79%	-3.36%	0.07
Wheat-MN	-0.57%	-0.25%	-0.32%	0.73	-4.33%	-3.15%	-1.18%	0.09
Rough Rice	4.32%	-1.93%	6.25%	0.73	-0.73%	-6.89%	6.16%	0.58
Soy-meal	-0.01%	-0.24%	0.23%	0.42	-1.92%	-2.87%	0.95%	0.27
Soy-oil	-1.42%	2.46%	-3.88%	0.03	-5.64%	3.02%	-8.66%	0.30
Oats	-4.40%	1.67%	-6.07%	0.01	-2.81%	-8.82%	6.01%	0.17
Live Cattle	-3.46%	-1.06%	-2.40%	0.05	-0.49%	2.37%	-2.86%	0.33
Lean Hog	-6.86%	-3.64%	-3.22%	0.00	0.43%	-1.65%	2.08%	0.83
Feed. Cattle	1.23%	3.61%	-2.38%	0.00	-1.57%	4.22%	-5.80%	0.25
Cocoa	4.02%	4.07%	-0.05%	0.00	2.86%	-5.00%	7.86%	0.10
Sugar	6.70%	-1.09%	7.79%	0.14	-6.76%	-0.24%	-6.53%	0.69
Coffee	2.55%	2.87%	-0.32%	0.14	-9.35%	-8.86%	-0.49%	0.00
Cotton	-0.18%	-2.46%	2.28%	0.21	-2.95%	-0.23%	-2.72%	0.67
Lumber	4.55%	-2.33%	6.88%	0.22	0.93%	-5.11%	6.04%	0.20
Orange Juice	1.52%	-0.77%	2.29%	0.17	-2.08%	-4.54%	2.46%	0.14
WTI	1.30%	-1.28%	2.58%	0.19	-5.61%	-4.51%	-1.10%	0.26

Table 3.5 (cont.)

Gold	-0.03%	-0.66%	0.63%	0.00	0.85%	2.73%	-1.88%	0.54
Copper	-1.11%	3.90%	-5.01%	0.01	-2.75%	4.19%	-6.94%	0.53
Silver	-0.80%	-0.61%	-0.19%	0.32	2.81%	3.53%	-0.72%	0.98
Average	0.28%	-0.01%	0.24%	0.15	-1.86%	-1.83%	-0.04%	0.36

Notes: This table reports the mean returns of the spread time series momentum (STSM) and long-only trading strategies for individual markets, and the return differences between the strategies. The 2nd (6th) column reports the annualized returns of the STSM-spread (STSM-outright) trading strategy. The long-only spread (outright) profitability is reported in the 3rd (7th) column. P-values measure statistical significance of return differences between the two strategies and are derived from the nonparametric Wilcoxon test. All returns are annualized.

Table 3.6: STSM and Long-Only Trading Strategies Return Comparisons (Portfolio)

Look-back (L)	STSM-Spread vs. Long-Spread						STSM-Outright vs. Long-Outright					
	STSM (Equal)	Diff.	p- value	STSM (Volat.)	Diff.	p- value	STSM (Equal)	Diff.	p- value	STSM (Volat.)	Diff.	p- value
1	0.39%	1.25%	0.47	0.72%	1.58%	0.25	-0.15%	-1.98%	0.28	1.22%	-0.61%	0.27
2	0.02%	0.88%	0.21	-3.75%	-2.89%	0.16	0.44%	-1.39%	0.25	-6.83%	-8.66%	0.37
3	-1.45%	-0.59%	0.28	-3.60%	-2.74%	0.25	-12.64%	-14.47%	0.55	2.66%	0.83%	0.28
4	0.30%	1.16%	0.30	0.14%	1.00%	0.28	0.13%	-1.70%	0.16	2.14%	0.31%	0.13
5	-1.19%	-0.33%	0.03	-0.70%	0.16%	0.10	-0.83%	-2.66%	0.14	1.74%	-0.09%	0.47
6	-0.13%	0.73%	0.41	-0.20%	0.66%	0.43	0.47%	-1.36%	0.12	0.70%	-1.13%	0.20
7	0.07%	0.93%	0.35	-2.45%	-1.59%	0.26	-4.58%	-6.41%	0.38	-5.42%	-7.25%	0.04
8	-2.19%	-1.33%	0.23	0.57%	1.43%	0.20	1.18%	-0.65%	0.13	1.05%	-0.78%	0.13
9	-0.28%	0.58%	0.28	-4.50%	-3.64%	0.04	-0.81%	-2.64%	0.23	1.32%	-0.51%	0.74
10	-1.18%	-0.32%	0.44	-0.52%	0.34%	0.38	3.11%	1.28%	0.32	-5.91%	-7.74%	0.11
11	-0.25%	0.61%	0.11	-3.88%	-3.02%	0.30	0.09%	-1.74%	0.26	1.06%	-0.77%	0.09
12	0.28%	1.04%	0.37	-1.47%	-0.61%	0.23	-1.86%	-3.69%	0.39	-1.47%	-3.30%	0.41
Mean	-0.49%	0.37%		-1.64%	-0.78%		-1.29%	-3.12%		-0.65%	-2.48%	

Notes: The table reports the annualized arithmetic mean portfolio returns of the spread time series momentum (STSM) strategy. Both the equal weight (equal) and volatility weighted (volat.) portfolio returns are reported with different look-back months (1-12). Columns 3 and 6 report the mean difference between the STSM-spread and the spread long-only strategies. Columns 9 and 12 report the mean difference between a STSM-outright and the long-only outright nearby contract strategy. Columns 4, 7, 10, and 13 report the p-values for the null hypothesis that monthly STSM and long-only returns are not statistically different from each other based on the nonparametric Wilcoxon test.

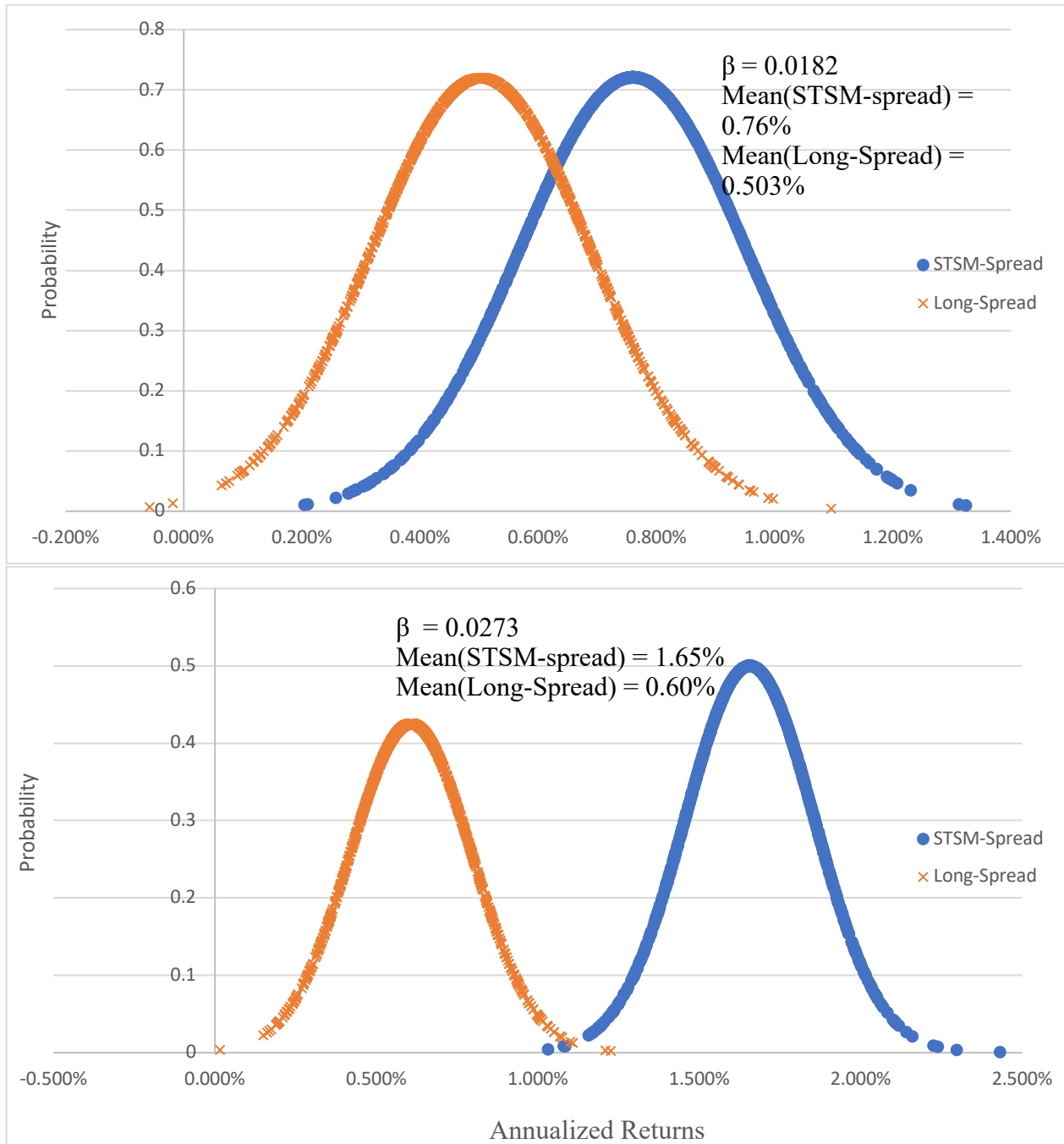


Figure 3.1: Distribution of Annualized Mean Returns from Bootstrapped Samples (STSM-spread and Long-only Spread Trading Strategies)

Notes: This figure shows the distribution of the annualized mean returns of a portfolio of the 22 commodities that adopts a STSM-spread strategy (blue color) and a portfolio that follows a long-only spread strategy (orange color). Data are simulated under two different predictability assumptions. We use 1,000 simulations that allow us to calculate 1,000 returns under each investment strategy. The imposed regression slope β is set to be 0.0182 (top panel) and 0.0273 (bottom panel). These values are two and three times larger than estimated predictability based on sample data, respectively.

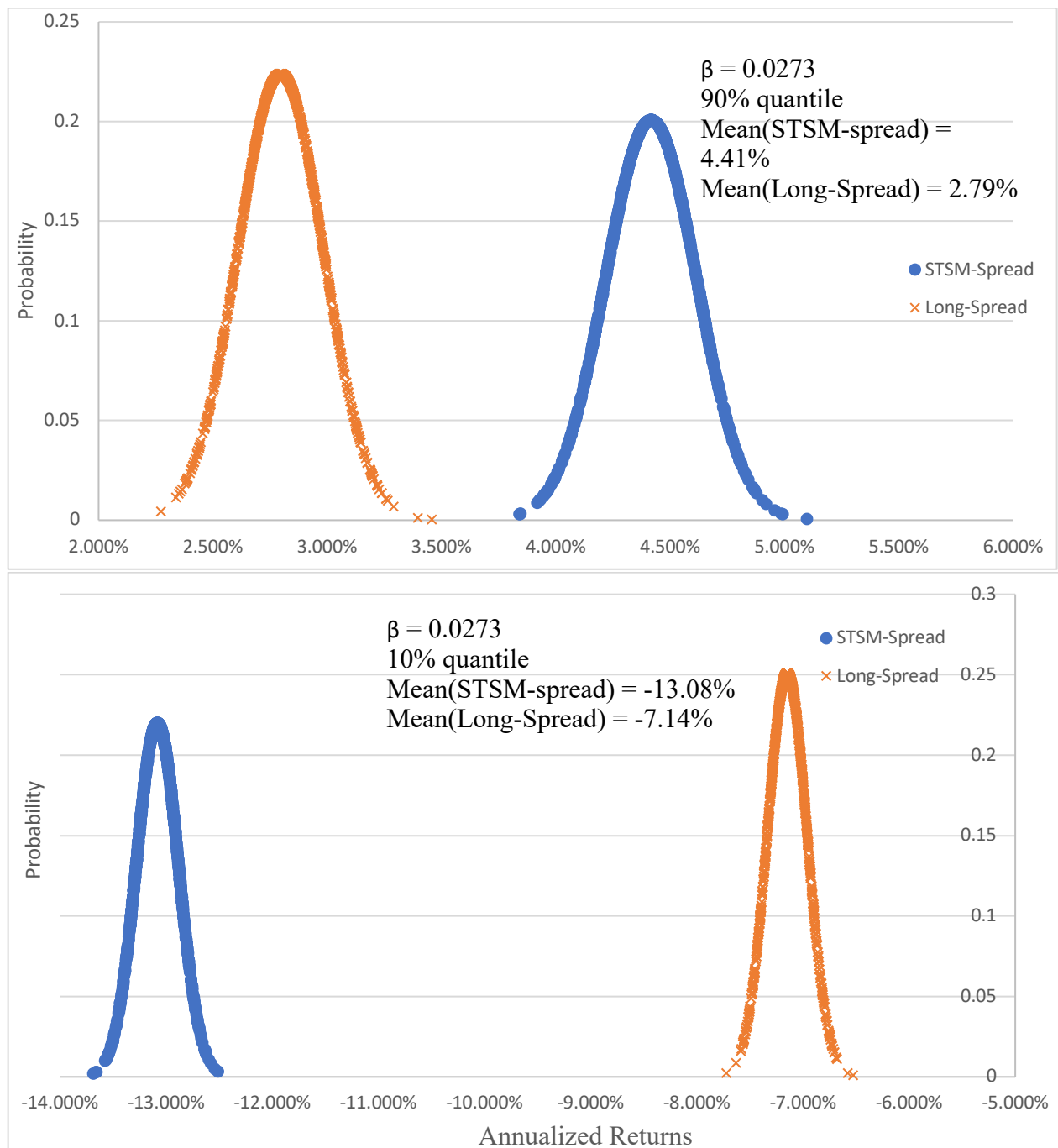


Figure 3.2: Distribution of 90 and 10 Percent Quantiles of Annualized Returns from Bootstrapped Samples (STSM-spread and Long-only Spread Trading Strategies)

Notes: This figure shows the distribution of 90% (top panel) and 10% (bottom panel) quantile of the annualized mean returns of the STSM-spread and long-only spread strategies. These are calculated for each of the 1,000 bootstrapped samples that impose a minimum degree of predictability. The regression slope β is set to be 0.0273 (three times larger than the estimated predictability based on sample data).

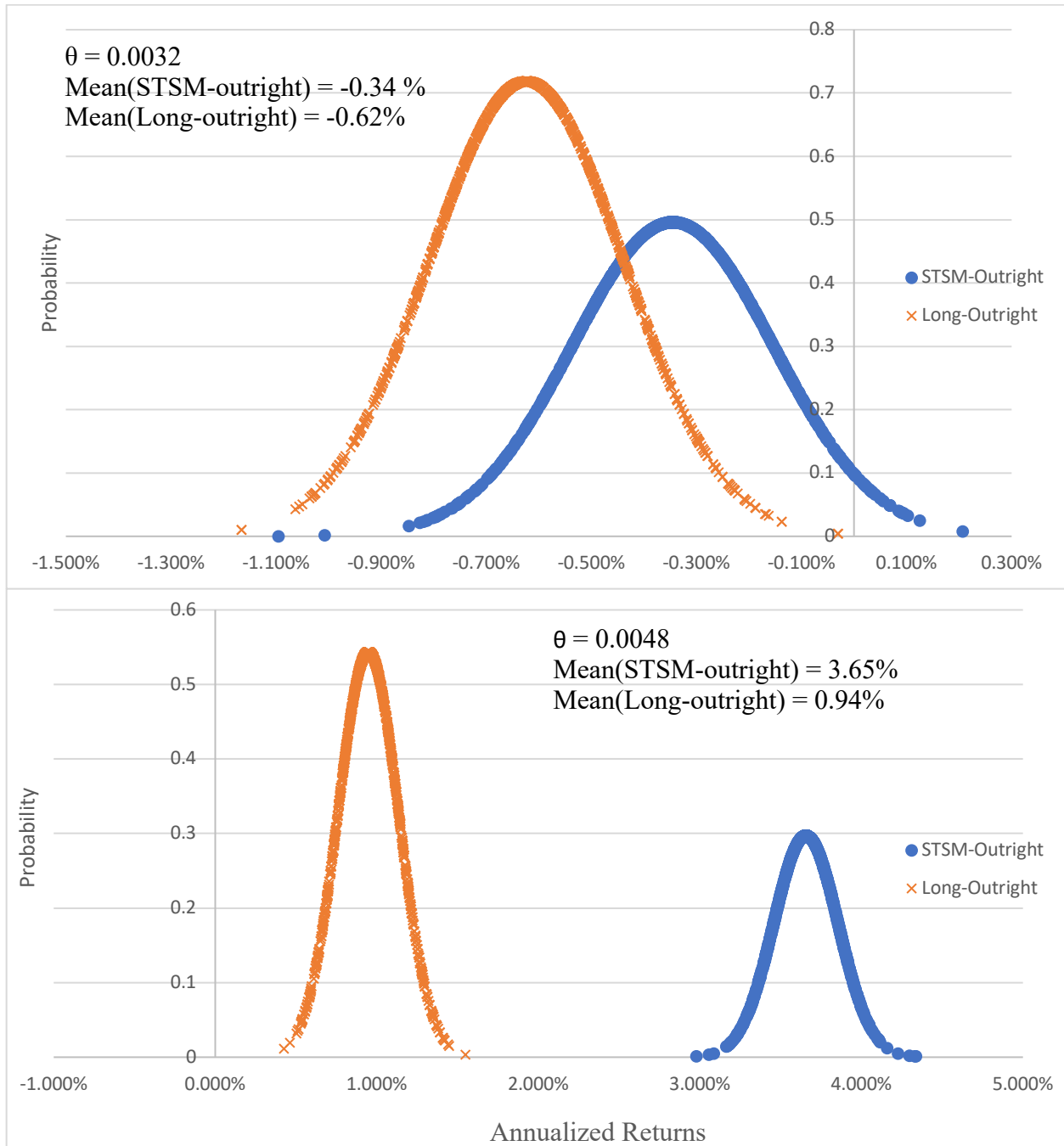


Figure 3.3: Distribution of Annualized Mean Returns from Bootstrapped Samples (STSM-outright and Long-only outright Trading Strategies)

Notes: This figure shows the distribution of the annualized mean returns of a portfolio of the 22 commodities that adopts a STSM-spread strategy (blue color) and a portfolio that follows a long-only spread strategy (orange color). Data are simulated under two different predictability assumptions. We use 1,000 simulations that allow us to calculate 1,000 returns under each investment strategy. The imposed regression slope θ is set to be 0.0032 (top panel) and 0.0048 (bottom panel). These values are two and three times larger than estimated predictability based on sample data, respectively.

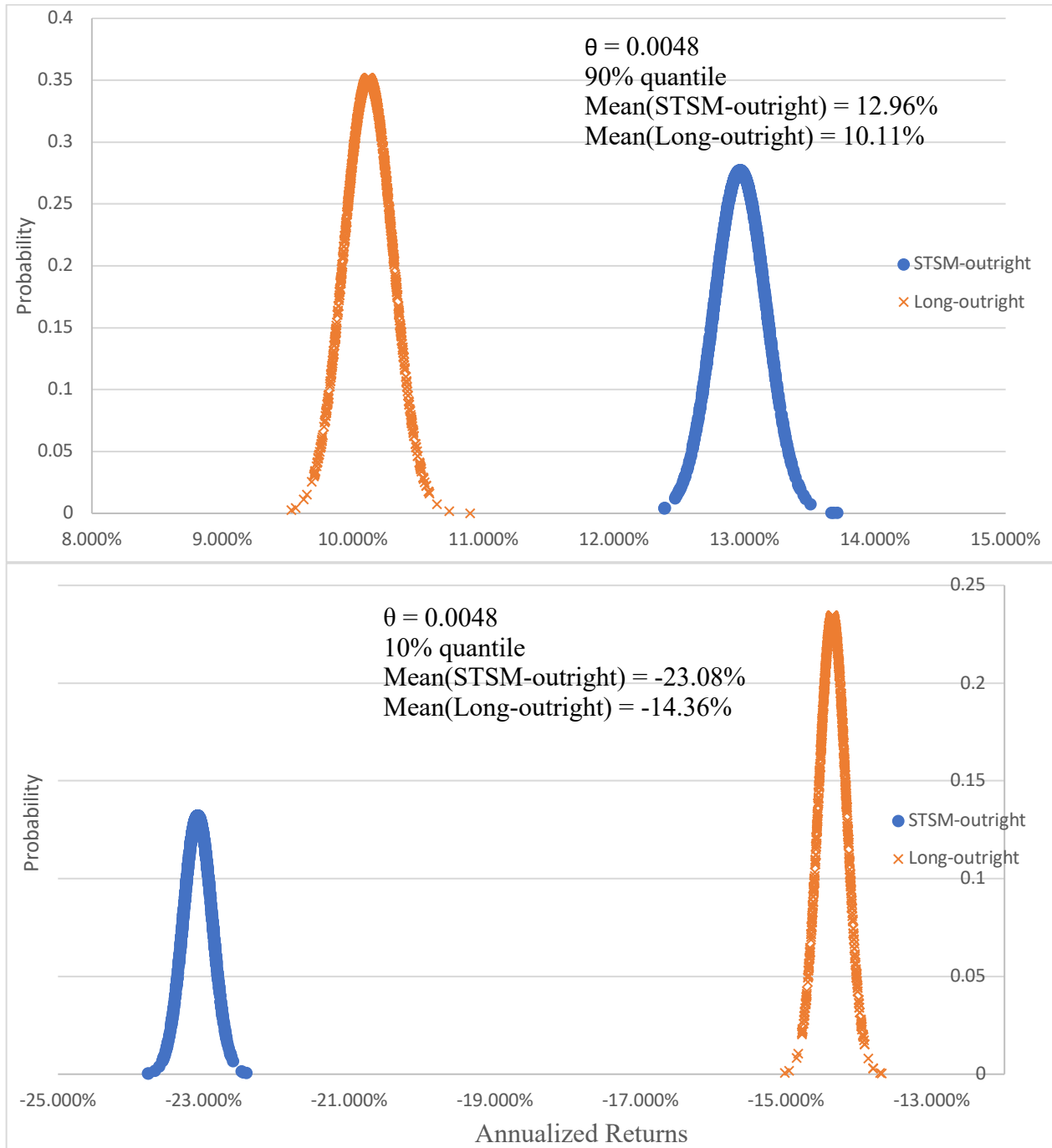


Figure 3.4: Distribution of 90 and 10 Percent Quantiles of Annualized Returns from Bootstrapped Samples (STSM-outright and Long-only outright Trading Strategies)

Notes: This figure shows the distribution of the 90% (top panel) and 10% (bottom panel) quantile of the annualized mean returns of the STSM-outright and long-only outright strategies. These are calculated for each of the 1,000 bootstrapped samples that impose a minimum degree of predictability. The regression slope θ is set to be 0.0144 (three times larger than the estimated predictability based on sample data).

3.7.References

- Ang, A. and Bekaert, G., 2007. Stock return predictability: Is it there? *The Review of Financial Studies*, 20(3), pp.651-707.
- Asness, C.S., Moskowitz, T.J. and Pedersen, L.H., 2013. Value and momentum everywhere. *The Journal of Finance*, 68(3), pp.929-985.
- Bessembinder, H., 2018. The “roll yield” myth. *Financial Analysts Journal*, 74(2), pp.41-53.
- Boons, M. and Prado, M.P., 2019. Basis-momentum. *The Journal of Finance*, 74(1), pp.239-279.
- Commodity Futures Trading Commission (CFTC). Strategic Plan: 2020-2024, 2020. Available at: https://www.cftc.gov/media/3871/CFTC202024_2024StrategicPlan/download
- Daniel, K. and Moskowitz, T.J., 2016. Momentum crashes. *Journal of Financial Economics*, 122(2), pp.221-247.
- Fama, E. 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, 25: 1043–53.
- Georgopoulou, A. and Wang, J., 2017. The trend is your friend: time series momentum strategies across equity and commodity markets. *Review of Finance* 21(4), pp. 1557–1592
- Gorton, G. and Rouwenhorst, K.G., 2006. Facts and fantasies about commodity futures. *Financial Analysts Journal*, 62(2), pp.47-68.
- Gorton, G.B., Hayashi, F. and Rouwenhorst, K.G., 2013. The fundamentals of commodity futures returns. *Review of Finance*, 17(1), pp.35-105.
- Hieronimus, T.A., 1977. Economics of futures trading for commercial and personal profit.
- Huang, D., Li, J., Wang, L. and Zhou, G., 2020. Time series momentum: Is it there? *Journal of Financial Economics*, 135(3), pp.774-794.
- Irwin, S.H. and Sanders, D.R., 2011. Index funds, financialization, and commodity futures

- markets. *Applied Economic Perspectives and Policy*, 33(1), pp.1-31.
- Irwin, S.H., Sanders, D.R., Smith, A. and Main, S., 2020. Returns to Investing in Commodity Futures: Separating the Wheat from the Chaff. *Applied Economic Perspectives and Policy* 42(4), pp. 583-610.
- Jegadeesh, N. and Titma, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, 48(1), pp.65-91.
- Kang, W., Rouwenhorst, K.G. and Tang, K., 2020. A tale of two premiums: The role of hedgers and speculators in commodity futures markets. *The Journal of Finance*. 75(1), pp.377-417.
- Kreiss, J.P. and Lahiri, S.N., 2012. Bootstrap methods for time series. *Handbook of statistics* 30, pp. 3-26. Elsevier.
- Kim, A., Tse, Y., Wald, J., 2016. Time series momentum and volatility scaling. *Journal of Financial Markets* 30, pp. 102–124.
- Lesmond, D.A., Schill, M.J. and Zhou, C., 2004. The illusory nature of momentum profits. *Journal of financial economics*, 71(2), pp.349-380.
- Leuthold, R.M., Garcia, P. and Lu, R., 1994. The returns and forecasting ability of large traders in the frozen pork bellies futures market. *Journal of Business*, pp.459-473.
- Miffre, J. and Rallis, G., 2007. Momentum strategies in commodity futures markets. *Journal of Banking & Finance*, 31(6), pp.1863-1886.
- Moran, N.M., Irwin, S.H. and Garcia, P., 2020. Who Wins and Who Loses? Trader Returns and Risk Premiums in Agricultural Futures Markets. *Applied Economic Perspectives and Policy*, 42(4), pp.611-652.
- Moskowitz, T.J., Ooi, Y.H. and Pedersen, L.H., 2012. Time series momentum. *Journal of Financial Economics*, 104(2), pp.228-250.

- Newey, W.K. and West, K.D., 1994. Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies*, 61(4), pp.631-653.
- Park, C.H. and Irwin, S.H., 2007. What do we know about the profitability of technical analysis?. *Journal of Economic Surveys*, 21(4), pp.786-826.
- Paschke, R., Prokopczuk, M. and Simen, C.W., 2020. Curve momentum. *Journal of Banking & Finance*, 113, p.105718.
- Sanders, D.R. and S.H. Irwin. 2012. A reappraisal of investing in commodity futures markets. *Applied Economic Perspectives and Policy* 34 (3), 515-530
- Singleton, K.J. , 2014. Investor flows and the 2008 boom/bust in oil prices. *Management Science* 60 (2), 300–318
- Szymanowska, M., De Roon, F., Nijman, T. and Van Den Goorbergh, R., 2014. An anatomy of commodity futures risk premia. *The Journal of Finance*, 69(1), pp.453-482.
- Tomek, W.G., 1997. Commodity futures prices as forecasts. *Applied Economic Perspectives and Policy*, 19(1), pp.23-44.
- Wang, Y., Liu, L. and Wu, C., 2020. Forecasting commodity prices out-of-sample: Can technical indicators help?. *International Journal of Forecasting*, 36(2), pp.666-683.
- Willenbrock, S., 2011. Diversification return, portfolio rebalancing, and the commodity return puzzle. *Financial Analysts Journal*, 67(4), pp.42-49.
- Working, H. (1970) Economic functions of futures markets. *Futures Trading in Livestock—Origins and Concepts*, Chicago: Chicago Mercantile Exchange, 267–297.
- Zakamulin, V. and Giner, J., 2020. Time Series Momentum in the US Stock Market: Empirical Evidence and Theoretical Implications. Available at SSRN 3585714.

3.8. Supplementary Results

3.8.1. Spread Momentum Predictability (with Roll Yield)

Table 3.7: Spread Momentum Predictability (1986-2019, 12-Month Look-back Period with Roll Yield)

	<i>STSM-spread</i>			<i>STSM-outright</i>		
	Slope	t-stat. (slope)	R ²	Slope	t-stat. (slope)	R ²
Corn	0.35**	1.93	0.028	1.18**	2.49	0.102
Soybeans	-0.09	-1.39	0.091	0.99	0.08	0.057
Wheat-KC	0.72***	3.47	-0.007	1.42**	2.17	0.009
Wheat-Chi	0.22***	4.79	0.182	0.72	0.79	0.075
Wheat-MN	0.62	0.05	0.301	3.07	0.55	0.105
Rough Rice	1.03	1.33	0	0.95***	3.17	0.021
Soy-meal	0.20**	1.93	0.077	1.65	1.03	0.012
Soy-oil	-0.13	-3.80	0.021	3.51	0.75	-0.081
Oats	1.35***	4.21	0.287	0.98**	2.19	0.135
Live Cattle	0.72***	3.77	0.133	1.03***	2.99	0.09
Lean Hog	0.62**	2.14	0.315	0.92	1.11	0.09
Feed. Cattle	0.35***	2.99	0.07	1.11***	2.91	0.093
Cocoa	1.26	0.56	0.224	-0.24	-0.48	-0.012
Sugar	-0.11	-2.87	0.133	0.38***	2.91	0.045
Coffee	0.13***	4.76	0.119	-0.62	-2.00	0.042
Cotton	0.42	0.24	0.28	3.69***	4.68	0.06
Lumber	0.61**	2.54	-0.014	1.55**	2.56	0.108
Orange Juice	1.72	0.81	0.259	3.91	0.94	0.048
WTI	0.30	0.68	0	1.09	1.19	-0.003
Gold	0.09***	4.28	0.266	-0.12	-2.65	0.033
Copper	0.52	1.56	0.063	2.62	1.59	0.018
Silver	-0.03	-0.74	0.231	1.00**	2.62	0.081
Average	0.38	2.99	0.139	0.92	1.58	0.051

Notes: This table reports results of the market-level regression analysis that investigates the ability of the spread momentum indicator (SM_t^i) to predict future spread (outright) returns (equation (3.6) in the paper). The dependent variable is the spread (outright) return, $R_{t+1}^{long-spread,i}$ ($R_{t+1}^{long-outright,i}$) for month $t+1$. The independent variable is SM_t^i calculated over the past 12-months as in equation (3.3) in the paper. The non-realizable roll yields are included in the returns' calculation. There are T=408 (12 months $\times$ 34 years) observations for each market. Column 2 (5) shows the regression slope for spread (outright) return predictability. The corresponding t-statistics

Table 3.7 (cont.)

in column 3 (6) are based on Newey and West (1994) standard errors to control for non-*iid* errors. In applying Newey-West, we use the automatic lag selection technique formula ($\text{floor}[4(T/100)^{2/9}]$). In our case, the lag for each individual market regression is 5. ***, **, * indicate significance at the 1%, 5% and 10% levels.

3.8.2. STSM Return (with Roll Yield)

Table 3.8: STSM Momentum Trading Strategies Annualized Return (With Roll Yield, Individual Markets)

	STSM-spread	STSM-outright
Corn	4.23%	13.48%
Soybeans	7.80%	-1.16%
Wheat-KC	10.84%	2.69%
Wheat-Chi	3.38%	-3.44%
Wheat-MN	5.86%	10.28%
Rough Rice	12.35%	0.16%
Soy-Meal	13.67%	-2.53%
Soy-oil	9.20%	-1.27%
Oats	1.88%	10.26%
Live Cattle	0.70%	5.19%
Lean Hog	5.90%	12.22%
Feed. Cattle	7.14%	8.44%
Cocoa	-3.11%	10.75%
Sugar	-6.63%	8.24%
Coffee	-0.09%	1.31%
Cotton	11.74%	7.90%
Lumber	4.39%	8.65%
Orange Juice	11.76%	10.68%
WTI	5.24%	-0.54%
Gold	5.99%	-3.41%
Copper	-0.61%	10.36%
Silver	10.65%	11.67%
Average	5.56%	5.45%

Notes: This table reports the mean annualized returns of the STSM trading strategies for individual markets. The returns include the non-realizable roll yields.

3.8.3. Spread Momentum Portfolio Returns (with Roll Yield)

Table 3.9: STSM Portfolio Returns (With Roll Yield)

Look-back	STSM-spread		STSM-outright	
	Equal Weight	Volatility Adjusted	Equal Weight	Volatility Adjusted
1	5.08%	6.89%	-2.08%	5.76%
2	9.59%	-0.12%	2.23%	0.02%
3	4.86%	12.13%	-0.95%	-1.55%
4	11.04%	12.12%	9.34%	14.91%
5	3.65%	-0.25%	14.37%	0.31%
6	4.86%	6.86%	0.93%	-5.19%
7	12.33%	10.59%	6.90%	9.88%
8	1.95%	11.10%	2.07%	8.57%
9	5.13%	8.73%	-2.91%	2.66%
10	4.58%	7.35%	0.33%	3.74%
11	3.93%	10.54%	12.87%	-1.60%
12	6.56%	10.68%	-1.09%	6.99%
Average	6.13%	8.05%	3.50%	3.71%

Notes: The table reports the annualized portfolio returns of the STSM trading strategies. Both the equal weight (equal) and volatility weighted (volat.) portfolio returns are reported with different look-back months (1-12). The returns reported here include the contract roll yield.

CHAPTER 4

THE ROLE OF SWAP CONTRACTS IN THE U.S. AGRICULTURAL COMMODITY MARKETS

4.1. Introduction

Agricultural commodity markets have experienced unprecedented price volatility over the past two decades placing risk management at the center of agribusinesses' operations. Financial derivatives such as futures, options, and other forward contracts are widely used price risk management tools (Prager et al. 2020) that have received different degrees of attention from the literature. Current knowledge on other forward contracts is scarce. Among the latter, swap contracts have become increasingly relevant in commodity markets (Irwin and Sanders 2011; Peterson, 2014; Mixon, Onur, and Riggs 2018), with their flexibility being an important factor behind their rapid adoption (Peterson 2018). Research on swap contracts has however been limited due to the opacity that characterized this market (Mixon, Onur, and Riggs, 2018).

Swaps remained relatively unregulated until the enactment of the Dodd-Frank Act in 2010. To promote market transparency, the Act mandated information on swap transactions such as price and quantity be registered and publicly released in real-time through designated swap data repositories (SDR). Despite increased transparency and in contrast to the attention that swap markets have received from regulatory agencies, little is known about these markets not only in the commodity space, but also in the broader swaps market. Few published articles can be identified in non-commodity swaps. Loon and Zhong (2016) assess credit default swaps' (CDS) liquidity costs after the passage of the Dodd-Frank Act. They conclude that the CDS market became more liquid with an average effective spread of only 0.27% of the asset price. As another layer of transparency, the Dodd-Frank Act promoted swap trading on Swap Execution Facilities

(SEFs) which are electronic platforms that serve as counterparty matching services (CFTC 2022). With proprietary data from swap clearing houses, Benos et al. (2020) show that interest rate swap markets' liquidity improved and transactions costs decreased after the introduction of SEFs. Riggs et al. (2020) use the quote and trade data from major SEFs for CDS and shed light on customers' trading strategies and dealers' liquidity provision. Consistent with earlier studies, they suggest that CDS' transactions costs are rather low.

Research on agricultural commodity swap markets is equally scarce. A series of studies use swap dealer positions from the CFTC weekly Commitment of Traders reports and focus on whether speculative activity drives futures prices through swaps. They find no convincing evidence that swap dealers drive futures market prices (e.g. Stoll and Whaley 2009; Sanders and Irwin 2010; Aulerich, Irwin, and Garcia 2010). Based on a confidential CFTC dataset, Mixon, Onur, and Riggs (2018) show that swap dealers hold relatively risk-neutral portfolios after aggregating their swap and futures contracts positions. This suggests that most (if not all) dealers' swap positions are ultimately hedged through futures. For example, if a hedger initiates a long (short) swap contract with a swap dealer, the dealer will hedge the risk by taking a long (short) position in the futures market (CFTC 2008; Mixon and Onur 2020). The role of swap dealers ensures a strong relationship between futures and swap markets and their prices, which in turn should warrant some degree of effectiveness of swaps as risk hedging instruments. Maples, Brorsen, and Etienne (2019) examine the hedging performance of swap contracts in fertilizer markets using the weekly Fertilizer Index as a proxy for fertilizer swaps' price. Their results suggest that swaps are effective risk management tools over relatively long hedging horizons.

Price discovery and risk transfer are two main functions of financial markets providing instruments for managing price risks (Working 1961; Garbade and Silber 1983). Through price

discovery, efficient markets generate prices reflecting traders' opinions on the value of the commodity and transmit this timely information along the marketing system. A commodity can be traded in multiple markets at the same time, either physically in the cash market or through financial derivatives. Since arbitrage ensures these markets transmit information to each other, the process of price discovery involves all these markets in different degrees, which can be captured through price discovery shares (Garbade and Silber 1983; Hasbrouck 1995; Yan and Zivot 2010). Price discovery is predicated on the existence of an equilibrium relationship among the markets involved and informs on the speed of different markets to incorporate new information. Markets with larger price discovery shares incorporate information faster than those with lower shares. To date, the analysis of price discovery has been primarily limited to cash and futures prices, where transparency and efficiency of futures markets usually allows them to dominate the price discovery process over cash markets. Within futures markets, price discovery has been studied across different exchanges when they trade related derivatives (Janzen and Adjemian 2017), contracts with different maturities (Hu et al. 2020), or different types of orders within the same contract (Arzandeh and Frank 2019). Efficient price discovery in commodity financial derivatives is at the core of hedging effectiveness: a financial instrument will allow a perfect hedge whenever its price and the cash price experience equal variations and thus incorporate the same information over the holding period (Ederington 1979; Garbade and Silber 1983).

The objective of this research is to shed light on the price discovery and hedging functions of corn and soybean swap contracts. We choose corn and soybeans as they are the top two crops planted in the U.S. and the most actively traded grains in the U.S. futures markets (Sanders, Irwin, and Merrin 2010). Our contributions to the literature are threefold. This is the first study that characterizes swap markets by providing descriptive statistics including volume, trade size,

intraday trading patterns, daily and seasonal trading patterns using the futures market as a benchmark for comparison. This helps us understand how the swap market compares in size and characteristics to the well-known futures market. Second, we shed light on price discovery. Despite the increasing relevance of swap markets, no study has investigated their role in price discovery in any market. Hence, our contribution goes beyond the agricultural economics literature. For the first time, we quantify the price discovery shares of the swap market against its futures counterpart by deriving Putnins (2013) information leadership share (ILS) using intraday data. In modern fast-moving electronic markets, high frequency data are essential to capture the speed of response of prices to different intraday events (He and Serra 2022).

Third, we test the hedging effectiveness of swap contracts in comparison with futures contracts. Maples, Brorsen, and Etienne (2019) is the only attempt to study the hedging effectiveness in commodity swaps. We improve on the latter by using swap contract price data rather than a proxy, comparing hedging effectiveness of swaps and futures, and using swap markets based on underlying assets (corn and soybeans) whose price is much more volatile than the fertilizer price used in Maples, Brorsen, and Etienne (2019). In our hedging analysis, we use daily prices which are lowest latency available for cash prices. This seems appropriate since earlier studies show that commercials tend to build long-term hedging positions in swap markets (Mixon, Onur, and Riggs 2018) and are less likely to be concerned about the vagaries of intra-day market activities. Our sample period is from January 2018 to December 2019, which encompasses a particularly turbulent period characterized by the U.S.-China trade war, and a series of poor weather events that delayed crop planting and increased price volatility. The relevant shocks to the fundamentals make our sample period especially suited to assess price discovery and hedging performance. We then compare the effectiveness of futures and swap contracts to hedge cash positions using the

Diebold-Mariano (1995) test. We provide insights on the minimum hedging horizons required for swap and futures contracts hedging to be equally effective.

Findings indicate that trade volume in swap markets represents less than 20% the trade volume in the futures market. We find that swaps and futures markets' prices are bonded by an equilibrium relationship and their price discovery shares are commensurate with their respective volume shares. Further, hedging in swaps is as effective as in futures for hedging horizons longer than 35 (50) days in the corn (soybean) markets. Swaps however are 3.8 to 5.1 times more expensive than futures contracts, which may reflect their flexibility to adapt to traders' needs.

4.2.Swap Markets

Although swap contracts are popular in interest rate and currency markets, their flexibility has made them increasingly relevant in commodity markets (Peterson, 2014; Mixon, Onur, and Riggs 2018; Peterson 2018). While there are no standardized swap trade conditions, the *plain vanilla* swap is the simplest and most popular form of swap contract (Carter 2015; Hull 2015). It works much like a cash-settled commodity futures contract; it does not involve physical delivery and is only settled financially (Cater 2015; Peterson 2018). One of the legs in a *plain vanilla* swap is held by the hedging counterparty and based on a fixed price agreed upon when entering the swap either through a short or long position. The other leg is usually based on a floating price calculated from a pre-determined benchmark price(s). At expiration, the hedging counterparty holding a short (long) position of a swap receives (pays) exactly the fixed price. The counterparty holding the floating price leg pays or receives from the hedging counterparty a cash flow that compensates the difference between the settlement price and the pre-agreed price.

Despite their commonalities with futures contracts, swap contracts differ in that they are highly customizable to counterparties' needs and traded with swap dealers over the counter (OTC). For

example, swap transactions can be written for any underlying commodity, with flexible notional amount and/or tenor. Parties to OTC swaps have greater freedom to bargain contract terms, including collateral and price, based on their creditworthiness. In addition, OTC swaps have lower margin requirements and less restrictive clearing mandates,¹⁶ which may be more convenient for those who face capital access constraints (Adjemian and Plato, 2010). Traders may also enter OTC swap contracts to avoid order management in derivatives markets. Earlier studies have demonstrated that commercials tend to place large swap orders with long tenors (Peterson 2014; Mixon, Onur, and Riggs 2018). If they were to place these orders in futures markets and to minimize transactions costs, commercials should build positions in the more liquid nearby contract and roll over the next deferred contract upon nearby contract expiration. The technical and managerial skills needed to conduct these operations through modern futures markets trading platforms are barriers to entry for medium or small sized commercials. OTC swap dealers are usually institutional traders who are more sophisticated in order management in derivatives markets and provide these traders the intermediate service to meet their hedging demand. Through their operations, swap dealers bridge the swaps and exchange-traded markets (CFTC, 2008; Mixon, Onur, Riggs, 2018; Mixon and Onur, 2021).

4.3. Data

We use corn and soybean SDR data for the period from January 3, 2018 to December 23, 2019. Corn and soybean futures prices and trade volumes are obtained from the Chicago Mercantile Exchange (CME) Market Depth (MD) database. Central Illinois daily cash market prices are

¹⁶ Based on a conversation with the Chicago Mercantile Exchange (CME) staff, there is no mandatory clearing mandate or standardized margin requirement for commodity swaps.

derived from the Commodity Research Bureau (CRB) database and represent the Farmers Grain Company of Central Illinois elevator price in Clinton, DeWitt County, IL. Details of the SDR and MD data are offered below.

4.3.1. Swap Data

The Dodd-Frank Act requires that all swaps (whether cleared or un-cleared) be reported to a registered swap data repository (SDR). There are three SDRs, the Depository Trust & Clearing Corporation (DTCC), the CME Swaps Repository, and the Intercontinental Exchange (ICE) Trade Vault. Mandatory reporting began as early as December 31, 2012, but agricultural commodity swap data were not published in real-time until February 28, 2013 (CFTC 2011, 2012, 2013; Nazareth and Rosenberg 2013). Regardless of execution venue, swap dealers must report their swap trades to one of the designated SDRs and all swap trades are subject to real-time reporting and dissemination to the public.¹⁷ Swap Execution Facilities (SEF) data are not available for commodities. However, these transactions are also subject to reporting mandates and their information is available through the SDR.

After executing a swap trade, dealers must report to a SDR the execution and dissemination date and time accurate to the nearest second “as soon as technologically practicable” (CFTC, 2011). Delayed dissemination is allowed for certain block trades, and pre-approval is required. For these trades, execution and dissemination timestamps differ. These represent less than 1% in our sample, with an average dissemination delay of 21.6 hours. We use the swap trade execution

¹⁷ The swap trade data can be viewed and downloaded from the following links: <https://www.cmegroup.com/market-data/repository/data.html#commodities>, <https://pddata.dtcc.com/gtr/cftc/dashboard.do>, and <https://www.icetradevault.com/tvus-ticker/#>.

timestamp in our empirical analysis since this should better reflect the linkages between swaps and futures markets.

Swap dealers also report effective date, end date, underlying asset name, price, notional amount and total value of the swap contract. The effective (end) date specifies the initial (last) day of the swap term, with the time between the two dates being referred to as the tenor of the swap contract. In this research, we consider all *plain vanilla* swap trades whose floating leg is settled based on either the CME corn or soybean outright futures contract prices. Prices in the corn and soybean swap trades are quoted in U.S. dollars per bushel. The size of swap trades is reported using the notional amount, in number of bushels of corn and soybeans. We eliminate any transaction with a missing trade price, notional amount, tenor or execution time stamp. Within the 392 trading days in our study, we observe 261,780 (161,093) CME corn (soybeans) swap transactions from the SDR data. After eliminating transactions with missing values, 237,073 (145,530) corn (soybeans) swap transactions remain in our sample.¹⁸

Swap transactions' size and price cannot be directly compared to those in futures markets. To facilitate comparisons, we convert swap volume into futures-equivalent contract size, which is 5,000 bushels/contract. Prices per bushel are not directly comparable between futures and swap markets due to the different temporal horizon of the two contracts. For each swap transaction i on day t , we transform the swap price ($P_{t,i}^S$) into a nearby futures-equivalent price, which we name the *swap implied nearby price* ($P_{t,i}^{sin}$). We use the cost-of-carry model for storable commodities

¹⁸ Among the three different SDRs, the DTCC SDR accounts for more than 85% of the total sample, which is similar to Peterson's (2014) findings.

(e.g. Pindyck, 2001; Main et al., 2018) to discount swap prices ($P_{t,i}^s$) into $P_{t,i}^{sin}$. The model can be expressed as

$$P_{t,i}^{sin} = P_{t,i}^s \times e^{C_{t,i}(a-s)}, \quad (4.1)$$

where a (s) is the number of days left on day t for the nearby future (swap) to expire. $C_{t,i}$ is the cost of carry associated to the swap transaction (t, i) calculated as follows. First we identify two futures contracts traded on day t with expiration dates $(t + a)$ and $(t + b)$ such that $(a < s < b)$.¹⁹ We measure the prices of these two futures contracts as the average best bid-ask quote midpoint during the 5-minute interval prior to the swap transaction (t, i) execution. The wide 5-minute window allows for imprecisions in swap reporting, as well as for the possibility that swap dealers may hedge their positions in the futures markets slightly before the reporting, to prevent revealing information that may harm their interests in the futures market. We denote these futures mid-quote prices as $\bar{F}_{t,i,mid}(t + a)$ and $\bar{F}_{t,i,mid}(t + b)$. The cost-of-carry $C_{t,i}$ per day associated to the swap transaction (t, i) can then be expressed as:

$$C_{t,i} = \frac{1}{b-a} \times \ln \left[\frac{\bar{F}_{t,i,mid}(t+b)}{\bar{F}_{t,i,mid}(t+a)} \right]. \quad (4.2)$$

For example, if a swap trade i occurs at 9:00:00AM on day t , we approximate the cost of carry using the average quoted midpoints based on all quote revisions from 8:55:00AM to 9:00:00AM for contracts with expiration dates $(t + a)$ and $(t + b)$. In our sample, 5.7% (4.1%) of corn (soybean) swap trade occurs when futures markets are closed. In this case, we use the $C_{t,i}$ derived from the previous sessions' last 5-minute trading interval. We also conduct robustness checks

¹⁹ If the swap contract end-date is earlier than the nearby futures contract expiration, we approximate $C_{t,i}$ based on the nearby and the first deferred futures contracts. In our sample, 7.2% (5.1%) of corn (soybean) swap transactions fall into this category.

based on a constant $C_{t,i}$ throughout each day. We calculate the daily constant carry cost using the futures settlement prices of the previous day $t-1$ ($\bar{F}_{t-1,i,\text{settlement}}(t+a)$ and $\bar{F}_{t-1,i,\text{settlement}}(t+b)$).

4.3.2. *Futures Market Data*

From the corn and soybean MD data we obtain futures transaction quantities and prices, and the top of the book quotes to derive the $p_{t,i}^{sin}$ as in equation (4.2). While MD contains more precise time-stamps, we round them to the nearest second in order to match swap and futures data. Our descriptive, price discovery, and hedging analyses are based on the nearby futures contract for both corn and soybeans, as it is the most highly traded contract, which results in larger liquidity and lower transactions costs (Wang, Garcia, and Irwin 2014; Hu et al. 2020). We refer to the nearby contract as the closest to maturity contract and the first deferred contract as the second closest to maturity contract. The nearby contract is replaced by the first deferred contract on the first day of the nearby maturity month. The September corn futures contract and August and September soybean futures contracts are lightly traded (Smith 2005; Wang, Garcia, and Irwin 2014) and thus dropped from the analysis. This results in rolling the position from nearby to the first deferred, on average, every 3 (2.4) months in corn (soybeans).²⁰

4.4. **Descriptive Analysis of Corn and Soybean Swap Markets**

In this section we provide daily summary statistics for corn and soybean swap contracts relative to their futures counterparts. Grain futures market conditions are substantially different on USDA announcement and non-announcement days, with the former being characterized by increased volume and volatility (e.g. Wang, Garcia, and Irwin 2014; Huang et al. 2022). No research has

²⁰ These timings are based on using March, May, July, and December contracts in corn and January, March, May, July, and November contracts in soybeans.

investigated whether swap market conditions change on announcement days compared with other trading days and here we shed light on this question. Table 4.1 presents summary statistics separately for USDA announcement and regular trading days. Announcement days include days when the World Agricultural Supply and Demand Estimates (WASDE), Crop Production, Prospective plantings, Grain Stocks, and the Acreage reports are released. There are 34 announcement days during our sample period.

The daily average swap contract volume expressed in futures-equivalent number of contracts is 30,239 (20,790) in the corn (soybean) market. During USDA announcement days and relative to non-announcement days, swap trading volume increases to 30,965 (21,168), a 2.3% (1.8%) in the corn (soybean) market, but average increases are not statistically significant in either market. The corn (soybeans) swap daily volume represents on average 16.8% (19.9%) of the futures nearby contract volume and fluctuates between 6.8% and 39% (6.9% and 45.3%). On USDA announcement days the average trade ratio is 13.2% (17.2%), which is 3.6% (2.7%) lower than the non-announcement trading days for corn (soybeans). These declines are statistically significant at least at the 10% level in both markets, which suggests that on announcement days futures volume increases proportionally more than swaps volume. This can be attributed to several factors. First, the release of public information offers profit opportunities for market participants. Prices in modern electronic markets adjust to new announcement information in a matter of minutes, if not seconds (e.g. Huang, et al. 2022) and are likely to experience relevant jumps (Couleau, Serra, and Garcia 2020). OTC swap trades rely on dealers to facilitate the trading process, which is less automated and more time-consuming than futures trading. This makes swap markets less suited to profit from very short-run price movements.

Overall, our swap-to-futures volume ratios are much larger than Peterson's (2014) findings that swaps only account for less than 1% of the futures volume. Peterson (2014) covers the period immediately after the inception of agricultural swap trade reporting, when data quality was a major concern (e.g. Miedema 2014; Ackerman 2013; Ackerman and Burne 2014; Peterson 2014). Since then, CFTC has made a series of improvements on the swap reporting system that may explain the differences (CFTC 2015, 2016, 2018). Also, the increase in swap trade over time may be an indication of improved market liquidity and more sophisticated swap dealer service (e.g. Wang, Garcia, and Irwin 2014; Mixon, Onur, and Riggs 2018; Mixon and Onur 2020; Peterson 2018; Maples, Brorsen, and Etienne 2019).

Although the swap market has a smaller daily volume compared to the futures market, the average swap trade size is much larger (table 4.1), on the order of 47 (54) futures-equivalent contracts for corn (soybeans). On USDA announcement days, the average corn (soybean) swap trade size is 57 (61) futures-equivalent contracts, 21.2% (12.9%) larger than on non-announcement days, with the increases being statistically significant. The average swap trade size reported in Peterson (2014) is around 36 (18) contracts in the corn (soybean) market. We find the average corn (soybean) swap trade size is about 6 (12) times larger than the average futures market trade size. The relative trade size does not significantly change during announcement days. In contrast to futures contracts that have standardized expiration dates, each swap contract has a different tenor. The average tenor for the corn (soybean) market is 126 (141) calendar days, or approximately 4-5 months. When USDA announcements occur, the average tenor for corn (soybean) swap contracts is 168 (175) calendar days. Table 4.3 in the supplementary results presents the percentile distribution of swap trade tenors.

Based on intraday data, we derive the realized volatility (RV) from the nearby futures (swap) returns calculated based on transactions (implied nearby) prices resampled every 5-minutes. This resampling allows our RV measures to be free from microstructure noise in the futures market (Hansen and Lunde 2006). With the exception of announcement days, swaps are more volatile than futures. The average RV in corn swap (futures) market is 5.979 (4.627), substantially below the RV in the soybean futures (swap) market of 10.303 (9.175). This may be related to the 2018 U.S.-China trade war which heavily impacted the soybean market but had a less relevant effect on the corn market. It may also be due to the per bushel soybean price being larger than the corn price (\$3.78/bushel and \$9.18/bushel, respectively) over the sample period. While the futures market's RV is significantly higher during announcement days, we only see a modest and statistically insignificant increase in swap market's RV during these days.

4.4.1. Swap Trading Volume Patterns

Figures 4.1 and 4.2 compare the daily trading volume (upper panel) and volume ratio (lower panel) in the swap and futures corn and soybean markets, respectively. In the upper panels, the left (right) y-axis represents the total number of contracts for the swap (futures) market. The volume correlation between the two markets is 0.505 (0.477) for corn (soybeans), a moderate correlation indicating that volumes in the two markets move in the same direction. From the plot, we observe several volume spikes both in corn and soybean swaps and futures. In particular, a volume spike around late March and early April 2018 is apparent and likely related to the U.S.-China trade disruptions that started in March, 2018 (Swanson, Coppess, and Schnitkey 2018). Despite the trade war having a larger impact on the soybean market than on the corn market, spikes are relevant in both markets and involve volumes twice as large as average volumes. Another cluster of spikes emerges around mid-2019 both in corn swaps and futures likely reflecting the impact of excess

rainfalls and planting delays in 2019, which caused heightened market volatility and trading volume (Irwin and Hubbs 2019). The daily swap to futures volume ratio bounces mostly around 17% (20%) in the corn (soybean) market. The corn swap-futures volume ratio spikes in July-2019, possibly due to delayed plantings. The soybean swap-futures volume ratio experiences a spike in late 2019, which may result from uncertainty of export demand given the U.S.-China trade war.

Figure 4.3 presents the average corn and soybean daily swap (left y-axis) and futures (right y-axis) volume patterns by month which follow each other closely. Both markets experience higher volume during the first half of the year when planting takes place than in the second half. Corn swap (futures) volume peaks in May, reaching a daily average volume close to 49,613 (243,472) futures-equivalent contracts. The May peak may reflect market uncertainties on planting delays during spring 2019. For soybeans, the swap (futures) market tends to be most active in March, with an average of more than 22,000 (140,000) futures-equivalent contracts per day. As the planting season concludes in late spring, there is a steady decline in trading volume until the harvest season in both markets. The slight increase in August corn trading volume may be due to the information shock from the USDA announcements that changed market sentiment in August, 2019 (Hubbs 2019).²¹ The daily soybean swap market volume increases from August to November, which can be due to uncertainty of the export demand and the next planting season.

Our results (figures 4.1, 4.2, and 4.3) suggest that daily and monthly swap market volume closely follows the futures market volume. However, it is still unclear whether any intraday pattern

²¹ The August-2019 USDA crop production report delivered a much larger corn yield estimate than the market expected (Hubbs 2019). The corn futures price was limit down immediately after the news release on August 12th, 2019.

exists in swap trading and whether it resembles the well-known intraday futures volume pattern. Using transaction timestamps available from the SDR and MD data, we aggregate trading volume for each 2-hour trading interval and present results in figure 4.4 for both futures (right y-axis) and swap (left y-axis) markets, respectively. As is the case with previous plots, the two markets follow similar patterns. We observe the swap and futures volume peak during the time window 8:00AM – 9:59 AM coinciding with futures market opening and decline quickly afterwards. Despite similarities, the portion of transactions that occur during the futures market evening session is much larger in the swap than in the futures market. Our results suggest that more than 80% of corn and soybean futures volume occurs during daytime session. However, only 62% (58%) of total corn (soybean) swap trade occurs during the futures market daytime session. Hence, a significant proportion of swap trades occur outside the futures market day trading hours. In particular, swap volume experiences a local maximum during the interval 18:00 – 19:59 coinciding with the opening of the CME corn and soybean futures evening trading sessions (at 19:00). Lehecka, Wang, and Garcia (2014) show that futures market volume and volatilities may extend into the opening of the evening sessions as traders adjust to new information from daytime. This market activity may facilitate swap dealers hedging their positions in the futures market for swap contracts executed late in the day. Additionally, this time interval corresponds to the Asian markets' openings. This 2-hour interval registers an average of 3,479 futures-equivalent contracts traded in the soybean swap markets, which is 36% higher than the corn swap volume in the same interval. Swap contracts allow foreign investors to gain risk exposure in the U.S. futures markets through a swap dealer' intermediary service without handling cross-border capital flows (Carter 2015). Asian countries are major importers of US soybeans, which might be the reason why soybean swaps volume overtakes corn during this time interval (CME 2021).

4.5. Swap-futures Price Discovery

In the absence of trading constraints, there should be free arbitrage between swaps and futures markets. Also, as discussed, swap dealers tend to hedge their positions in the futures market that is most related to the underlying asset in swap contracts (CFTC 2008; Adjemian and Plato 2010; Carter 2015; Mixon and Onur 2020). By bridging the two markets, swap dealers strengthen the information transmission between them, which should result in the two prices following the same stochastic trend and thus being bonded by an equilibrium relationship. We quantify the relative contribution of each market on the commodity price discovery.

The literature on price discovery is extensive. Garbade and Silber (1983) proposed to derive price discovery shares of cash and futures markets based on a vector autoregression model. More recent price discovery models are developed in the framework of vector error correction models (VECM) which capture both long and short-term price linkages (e.g. Hasbrouck 1995; Yan and Zivot 2010; Putnins 2013). These include Baillie et al. (2002)'s component share (CS) and Hasbrouck (1995)'s information share (IS) price discovery measures. These measures, however, are reliable only between two price series with similar level of noise (Yan and Zivot 2010; Putnins 2013). From our earlier results in table 4.1, futures and swaps trading volume and frequencies are largely different from each other. The futures market is characterized by small orders, large trading volume and high trading frequencies which results in noisy price discovery. In contrast, the swap market experiences larger trade size, and much lower trading frequency and volume, which is less likely to generate noise. The IS and CS measures tend to overestimate the price discovery shares of the less noisy market. In our study, we use the information leadership shares (ILS) from Putnins (2013) which provides a more robust price discovery share under different levels of market noise. We present the method used to derive the ILS in the supplementary section 4.10.3.

We derive the ILS measures for each day in our sample. To do so, we estimate a reduced-form VECM using the intraday corn and soybean futures and swap-implied log prices. This requires pairing swap and futures price data. On average, we observe one corn (soybean) swap transaction every 2.23 (3.74) minutes. This contrasts with one corn (soybean) trade every 0.03 (0.07) minutes in the futures market day trading session, and every 0.33 (0.64) minutes during the evening sessions. Also, as discussed, the literature shows that microstructure noise could last as long as five minutes in the electronically traded security markets (Hansen and Lunde 2006). To minimize the impacts of market microstructure noise while being able to capture which price, futures or swaps, moves first we resample our intraday data into 5-minute level frequency by taking the last transaction price from each interval.²² When there is no transaction inside an interval, we replace the missing value with the most recent past transaction price in the market.

Following previous studies (Janzen and Adjemian 2017; Hu et al. 2020), we include a constant in the cointegration relationship to capture potentially permanent differences between prices within each day. We select the lag length for each daily VECM model using the Akaike Information Criteria (AIC), which chooses between two to seven lags for both corn and soybean models. The results from the Bayesian Information Criterion (BIC) are largely the same as AIC. Similar to Hu

²² Other price discovery studies adopt the event time resampling technique instead of clock time resampling. The event time resampling requires a “match” between price series from two markets, so that the two series have the same number of observations. This results in the less traded market having rather stale prices. This technique tends to work more effectively between markets with similar trading patterns (Hasbrouck 2021). Given the less-frequent-trading nature of the swap markets, we choose the clock time resampling strategy in our study.

et al. (2020), we use the Johansen rank test to examine the cointegration between intraday futures and swap prices. Out of the 392 days in our sample, we find 357 (372) days are characterized by a cointegration rank of 1 in the corn (soybeans) market. Most remaining days are characterized by a cointegration rank of 2, which is only possible when all prices are stationary. Since the first order difference of a stationary series is still stationary (Engle and Granger 1987), and VECM can be fit to both nonstationary and stationary data (Keele and De Boef 2004), we do not change price discovery methods for those cases. The days when prices are not cointegrated are excluded from the price discovery analysis.

Earlier results in our study show that swap markets experience different patterns between day and evening trading hours and regular and USDA announcement days. To provide more insights on the changes of price discovery shares across different market conditions, we report the ILS shares separately for USDA announcement and non-announcement days and day and evening trading hours. Table 4.2 reports the average information shares for swap and futures markets. Since the ILS of futures and swap markets add to 1, ILS measures can be interpreted as proportions. These measures suggest that the corn futures market contributes on average 82.9% to price discovery, while the swaps market contributes 17.1% during regular day-trading sessions. The swap price discovery share (17.1%) is relatively close to its volume share (16.9%), which is consistent with Hu et al. (2020) who suggest that volume share is a strong driver of information share. During USDA announcement days, the corn swap market information (volume) share drops by 4.9% (4.7%) compared with regular trading days. We also observe a similar pattern in the soybean market. During regular trading days, the soybean swap market contributes 18.3% to the price discovery share, which is 1.8% higher than on USDA announcement days. This is consistent with the soybean swap volume share of 19.1% (17.7%) during regular (announcement) trading

days. As discussed, swaps are less appropriate tools to profit from very fast price changes on announcement days, which contributes to their reduced relevance in both price discovery and volume shares. Also, swap market liquidity provision is largely dependent on swap dealers' risk preferences (Mixon and Onur 2020), who may be more risk averse amid high price volatilities during announcement days. This may have a negative impact on swap intermediation services provided by swap dealers and contribute to the decline in the price discovery share and trading volume ratio.

In contrast to day trading sessions, price discovery shares during evening trading hours are much larger in both corn and soybean swap markets. The corn (soybean) swap market price discovery share is 32.1% (40.7%) during evening trading hours, which is 15% (22.4%) higher than the day trading hours. At the same time, the swap volume ratio also increases dramatically during evening trading hours, specially for soybeans. The relatively large evening-session soybean swap price discovery and volume shares may be related to the previously discussed opening of Asian markets (CME 2021).

Despite the fact that the CFTC requires swap transactions to be reported “as soon as technologically practicable,” not all swap trades are executed in a SEF where relatively low-latency and fast reporting is feasible (Riggs et al. 2020). Swap trades executed off-facility may be processed manually and may experience reporting delays, which could raise questions on the accuracy of swap trade timestamps. We conduct a robustness check of our price discovery analysis by reprocessing (only) the swap trade timestamps to compensate for potential reporting delays. We move each swap trade timestamp by 1-, 3-, and 5-minute earlier than its original time (i.e. in

the 1-minute case, 8:30:00 becomes 8:29:00), re-compute the information shares²³ and present results in table 4.4 in the supplementary results section (4.10.2). Overall, we find the swap market price discovery shares grow gradually as the magnitude of the timestamp mismatch increases. This finding indicates that swap market price discovery shares may actually be larger if systematic reporting delays occur. Thus, our ILS estimation results from table 4.2 provide a conservative estimate (lower bound) of the swap market information shares.

4.6. Swap Contracts Hedging Effectiveness

In this section we study the performance of swap contracts relative to futures contracts in hedging corn and soybean positions in cash markets. Cash markets are only observed on a daily basis which requires to derive daily swaps and futures prices. In contrast to futures, swaps are not traded in transparent exchanges and are not mark-to-market securities with a daily settlement price. Instead, we compute a daily volume-weighted average swap implied nearby price as $P_t^{sin} = \sum_{i=1}^n P_{t,i}^{sin} \times \frac{Volume_{t,i}}{TotalVolume_t}$. Where $Volume_{t,i}$ is the corresponding swap trade i size on day t , measured in futures-equivalent contracts, and $TotalVolume_t$ is the total swap trading volume on day t . Similarly, we also compute the daily volume-weighted average futures nearby price. These prices along with the daily cash price are presented in figure 4.5 which suggests cash and derivatives prices closely follow each other. Also, futures and swap prices are tightly intertwined, which is consistent with our price

²³ With the reprocessed time stamps, we recalculate $P_{t,i}^{sin}$ from equation 4.1 and 4.2. As a robustness check, we also examine the price discovery shares with the swap-implied prices derived from a constant daily carrying cost as described in section 4.3. Results are nearly identical to what we present here, which is likely due to minimal intraday carrying cost changes.

discovery results and previous studies claiming that futures prices are often used as references for swap contract pricing (Adjemian and Plato 2010; Mixon and Onur 2021).

Earlier studies claim that swap markets register higher transactions costs than futures markets (Mixon, Onur, and Riggs 2018; Maples, Brorsen, and Etienne 2019). This should result in a buyer (seller)-initiated swap implied nearby price to be higher (lower) than the futures market quote midpoint. As a proxy for swap transactions costs (STC), we calculate the absolute relative difference between each swap price ($P_{t,i}^{sin}$) and the average nearby futures quote midpoint

($\bar{F}_{t,i,mid}(nearby)$) as $STC_{t,i} = \frac{|P_{t,i}^{sin} - \bar{F}_{t,i,mid}(nearby)|}{P_{t,i}^{sin}}$. Similar to our previous analysis, we measure

$\bar{F}_{t,i,mid}(nearby)$ during the 5-minute interval prior to the swap transaction (t,i) execution. We then

calculate the daily volume-weighted $STC_t = \sum_{i=1}^n STC_{t,i} \times \frac{Volume_{t,i}}{TotalVolume_t}$, which averages 0.358%

(0.499%) for corn (soybean). We use the percentage of effective spread as a proxy for futures

markets transactions costs (FTC). The formula for the effective spread for each transaction i on

day t is: $FTC_{t,i} = \frac{F_{t,i}(nearby) - F_{t,i,mid}(nearby)}{F_{t,i}(nearby)} \times 100$, where $F_{t,i}(nearby)$ is the nearby futures transaction

price of trade i on day t , $F_{t,i,mid}(nearby)$ is the prevailing quote midpoint value immediately before

each trade execution. We average $FTC_{t,i}$ ($\overline{FTC}_{t,i}$) using all transactions in our sample period, which

results in 0.069% (0.047%) $\overline{FTC}_{t,i}$ for corn (soybeans). Considering an average futures contract roll

period of 3 (2.4) months and 4 (5) month swap tenor in corn (soybeans) (see section 4.3.2 and table

4.1), we express the transaction costs in per-month-basis. These results suggest that swaps

transactions costs are about 3.8 to 5.1 times larger than futures transactions costs. Larger

transactions costs may be a reflection of the price that swap dealers charge to tailor contracts

according to customer needs (Maples, Brorsen, and Etienne 2019). Mixon and Onur (2020)

provide evidence that the provision of swap intermediation services by swap dealers is largely

dependent on their risk appetite. Transactions costs are variable and likely to be higher during

times of high volatility and low liquidity, which characterized our sample period. Soybeans were highly affected by the trade war, which may have resulted in specially large transactions costs in the soybean swap market.

Market participants in the commodity business can reduce risk exposure by taking a position in a derivatives market that offsets their cash positions. While the hedging effectiveness (HE) of futures markets has been thoroughly investigated (e.g. Ederington 1979; Myers and Thompson 1989; Wang, Wu, and Yang 2015), we know very little about swap contracts as hedging instruments. Mixon, Onur, and Riggs (2018) indicate that commercials in the energy sector have a larger footprint in the swaps than in the futures market. However, they do not examine the HE of swap contracts. As discussed, Maples, Brorsen, and Etienne (2019) investigate the HE in fertilizer swaps using a proxy for swap prices. They suggest that swaps provide risk protection to cash positions over relatively long hedging horizons. However, they do not compare the relative HE between futures and swap markets, which we study here using the daily prices described earlier.

A relevant hedging decision concerns selection of the hedging ratio (HR), which is the proportion of cash positions offset in the derivatives market. The naïve one-to-one hedging strategy consists of hedging the entire cash position, which results in a constant HR of 1. The naïve hedging strategy assumes that cash and futures prices move by the same amount during the hedging period and thus incorporate the same information. This implies that the derivatives market is an efficient forecast of the cash price. Extensive research has developed more sophisticated models to compute the optimal HR. However, Wang, Wu, and Yang (2015) demonstrate that sophisticated hedging models do not lead to performance that is statistically superior to the naïve hedging strategy. In our study, we use the naïve hedging strategy to compare the hedging performance of

futures and swap markets. Our goal here is not to select the hedging model that leads to most accurate hedging performance. Instead, we want to compare the swap and futures hedging performance under the same model.

A second issue in measuring HE concerns the choice of the hedging horizon, which is the time interval during which hedgers hold the position in the derivatives market. Earlier studies show that HE is sensitive to different horizons and tends to increase as the hedging horizon increases (e.g. see Wang, Wu, and Yang 2015 for the futures market and Maples, Brorsen, and Etienne 2019 for the swaps market). Over longer time horizons, if cash and futures/swaps prices are cointegrated, they will reflect the same information, which increases HE. In contrast, in the short-term cash and futures/swaps price changes are more susceptible to reflect noise and transitory components which can lead to lower HE. We examine a range of hedging horizons (τ) from 5 to 150 calendar days in increases of 5 days. The maximum hedging horizon in our analysis (150 days) covers the average swap tenor duration (table 4.1) for both corn and soybean markets. The conventional approach to measuring HE is to compare return variances between an unhedged (cash price only) and a hedged (cash and future/swaps) portfolio (Ederington 1979; Wang, Wu, and Yang 2015). To maximize the observations available to evaluate the HE, we follow Wang, Wu, and Yang (2015) and Maples, Brorsen, and Etienne (2019) and use a rolling window method with a fixed number of days equal to τ . Hence, we generate $(T - \tau + 1)$ portfolio returns over our sample containing a total of T days. The first return is generated from the window containing observations from day 1 to τ , the second rolling window encompasses observations from 2 to $\tau+1$ and so forth. We identify each rolling window using the end date of the window $t = \tau, \dots, T$.

4.6.1. Hedging Effectiveness (HE) Measure

Let $R_{cash,t}^\tau$, $R_{fut,t}^\tau$, and $R_{swap,t}^\tau$ denote the returns in the cash, futures, and swap markets obtained within the rolling window $t = \tau, \dots, T$ covering the hedging horizon (τ), respectively. Under naïve hedging, the return of the hedged portfolio at t is denoted as $(R_{cash,t}^\tau - R_{fut,t}^\tau)$ and $(R_{cash,t}^\tau - R_{swap,t}^\tau)$ if positions are taken in the futures and swap market, respectively. The HE quantifies the variance reduction of a hedged position over an unhedged position which is calculated using the returns across all rolling windows with the same hedging horizon as:

$$\begin{bmatrix} HE_{cash-fut}^\tau \\ HE_{cash-swap}^\tau \end{bmatrix} = \begin{bmatrix} 1 - \frac{Var(R_{cash,t}^\tau - R_{fut,t}^\tau)}{Var(R_{cash,t}^\tau)} \\ 1 - \frac{Var(R_{cash,t}^\tau - R_{swap,t}^\tau)}{Var(R_{cash,t}^\tau)} \end{bmatrix}. \quad (4.3)$$

$Var(\cdot)$ is the portfolio return variance across rolling windows. $HE_{cash-fut}^\tau$ ($HE_{cash-swap}^\tau$) measures the HE when using futures (swaps) to hedge against cash price risks. The closer the HE is to 1, the larger is the proportion of cash price risks that are offset by the financial instrument (futures or swaps) and thus, the more effective the hedging strategy. A $HE = 1$ (alternatively, $HE = 100\%$) refers to a perfect hedge where the hedged portfolio experiences zero return variances. In contrast, $HE = 0$ ($HE = 0\%$) suggests that the hedged portfolio has equal amount of risks as the cash position itself, implying that the financial instrument mitigates zero amount of cash price risks.²⁴

The HE only provides a point estimate of the hedging performance, which does not allow to compare the two HE measures in (4.3) statistically. To derive statistical inference we follow Wang, Wu, and Yang (2015) and conduct the Diebold and Mariano (DM 1995) test on portfolio squared

²⁴ It is possible for the HE to be negative, which indicates the hedged portfolio is riskier than the cash position alone.

returns. For every rolling window we define a loss function that represents the squared returns of each hedged portfolio as follows:

$$\begin{bmatrix} L_{cash-fut,t}^\tau \\ L_{cash-swap,t}^\tau \end{bmatrix} = \begin{bmatrix} (R_{cash,t}^\tau - R_{fut,t}^\tau)^2 \\ (R_{cash,t}^\tau - R_{swap,t}^\tau)^2 \end{bmatrix}. \quad (4.4)$$

where $L_{cash-fut/swap,t}^\tau$ is the loss function associated with each hedging horizon (τ), hedging instrument (*fut or swap*) and window (t) under the naïve hedging strategy. A larger loss suggests bigger return differences between cash and the financial instrument positions over τ , which leads to worse hedging performance. The loss differential between the cash-futures and cash-swaps portfolios is $D_t^\tau = L_{cash-fut,t}^\tau - L_{cash-swap,t}^\tau$. To control for potential serial correlation in the loss differentials, we calculate the DM statistic using the robust standard error estimator. Following Giacomini and White (2006) and Diebold (2015), we first run a regression of the loss differentials on a constant. The unconditional DM test is equivalent to testing whether the intercept is statistically different from zero, with a zero-intercept indicating no significant difference between the two hedging strategies. The Newey and West (1987) method is applied to correct heteroscedasticity and autocorrelation in the standard errors.²⁵

Figure 4.6 provides the HE results. The solid blue (orange) line measures the future (swap) contract HE, light blue bars the absolute value of the DM statistic and the silver dashed line the test statistic critical value. The upper (lower) plot shows the HE of swaps and futures in corn (soybeans). The HE improves as the hedging horizon grows. The corn futures (swaps) HE for a

²⁵ In applying Newey-West, we use the automatic lag selection formula: $(\text{floor}[4(T/100)^{2/9}])$.

Depending on the length of different hedging horizons, our lag length selection ranges from 4 to 5.

hedging horizon of 5 days is 27% (14%), which indicates that futures hedging mitigates 13% more risks than swap contracts. Our results are similar to Maples, Brorsen, and Etienne (2019) who find that swap HE is 20% or less during the one week hedging window in fertilizer markets. Corn swaps and futures HE improves steadily as the hedging horizon increases from 5 to 30 days.

While futures contracts outperform swaps initially, for hedging horizons beyond 35 days, both futures and swaps HEs converge to 80% and they are not statistically different from each other. Our swap HE measures over longer horizons are higher than those derived by Maples, Brorsen, and Etienne (2019), whose limited HEs are likely due to a weak cointegration between U.S. fertilizer cash prices and the port prices in New Orleans where the swap index is benchmarked. Similar to corn, the soybean futures HE is 17% higher than swaps under the hedging horizon of 5 days. The HE superiority of futures relative to swaps shrinks quickly from 5 to 30 days and ceases to be statistically significant for horizons 50 days or longer. Over short-hedging horizons, we find the soybean futures-swap HE performance gap to be larger than the corn market, which may be attributable to two reasons. First, relatively high transactions costs may result in low HE over short-horizons. Second, increased volatility related to the U.S.- China trade war may have contributed to larger basis risk and reduced HE (Adjemian, Smith, and He 2021).

Overall, our results suggest that swaps underperform the futures contracts for short hedging horizons. This may be due to differences in the speed of price discovery between the two markets. As discussed, futures hold the largest price discovery share, which means that new information is more often reflected in the futures market first, with swaps following subsequently. This may facilitate a better performance of futures in short-term hedging. Moreover, we calculate the HE based on transactions prices which include transactions costs. As shown earlier, transactions costs in swap contracts are likely higher than in futures contracts, which can lead to lower HE in the

short term where we assume rebalancing occurs fairly often through short-hedging windows. However, transactions costs may have far less impacts on long-hedging horizons. For hedging horizons longer than 35 (50) days in the corn (soybean) market, swaps and futures HE are analogous. Based on the swap tenor distribution from table 4.3 in the supplementary results, we find a vast majority of swap trades have tenors that are longer than the minimum threshold for effective swap hedging. Our findings are consistent with Mixon, Onur, and Riggs (2019) who argue that swap contracts are effective tools for commercial risk management.

4.7. Conclusion

Over the last couple of decades, swap contracts have become increasingly relevant in commodity markets which has been attributed to their flexibility relative to futures contracts (Irwin and Sanders 2011; Peterson 2014; Mixon, Onur, and Riggs 2018; Peterson 2018). Despite their increased relevance, they have received little scholarly attention resulting in important knowledge gaps on the relative magnitude and the role that swap markets play as a financial derivatives market. Using real-time released swap trade data from the Swap Data Repositories (SDR), we investigate corn and soybean swap markets from three different perspectives. We provide for the first time a general descriptive analysis on how these markets compare in size, trade volume, and patterns with their futures counterparts. We also shed light for the first time on their role in market price discovery. Third, improving on previous research, we further show swaps effectiveness in hedging cash positions relative to futures markets. Price discovery and hedging effectiveness of swap contracts are studied as they are two fundamental roles of financial derivatives markets (Working, 1961; Garbade and Silber, 1983).

We find evidence that swaps trade size is usually much larger (6 to 13 times) than the futures nearby trade size, though the overall swap market volume is only 17- 20% of the nearby futures

contract volume. Also, the two markets tend to have similar trading volume patterns both over different seasons and within each day. Swap and futures market prices closely track each other, which is consistent with swap dealers hedging their swap positions in the futures market around the same time the swap trade is executed. This ensures timely information transmission between the two markets. While futures and swap prices are linked by a long-run equilibrium, futures markets tend to incorporate new information faster at the intraday level. The price discovery shares between futures and swap markets are commensurate with their relative volume shares both in corn and soybean markets. However, our results also suggest that swap market price discovery shares increase substantially during evening trading hours, as their volume shares also increase relative to daytime sessions. The increase in both information and volume shares during evening trading sessions can be attributed to the active participation of foreign traders (Carter 2015; CME 2021), specially from Asia, in the U.S. swap markets.

The futures market is a popular risk management tool for cash positions. Our hedging effectiveness (HE) analysis shows that swaps are as effective in mitigating cash risks over long-hedging horizons as futures contracts. Futures hedging is superior to swaps for short-hedging horizons, which is consistent with price discovery in futures markets taking place faster than in swap markets. However, differences in HE between futures and swaps disappear for hedging horizons 35 (50) days or longer in the corn (soybean) market. Since a vast majority of swap trades have tenors beyond these horizons in our sample, our empirical findings suggest that swap contracts are effective hedging tools for commercials (Mixon, Onur, and Riggs 2018; Maples, Brorsen, and Etienne 2019).

Our results further suggest that swap contracts have larger transactions costs than futures contracts (3.8 to 5.1-fold). The largest transactions costs are registered in the soybeans market,

which needs to be understood in the framework of the U.S-China trade war and its large impacts on the soybean market. Our HE results and transactions costs estimates generate the question of why traders are willing to hedge through swaps at higher costs than futures. While our data do not allow us respond directly to this question, previous literature points at the flexibility of swap contracts over futures contracts, which allows small and mid-size commercials who face significant barriers to entry into the futures market to hedge their cash positions through swaps. Previous research also suggests that barriers to entry are likely related to the ability to compete in the fast-moving futures market, the minimum capital requirements to participate in futures markets and the risk of margin calls. Additionally, swap contracts also provide international traders a vehicle to hedge in U.S. markets without handling cross-border fund flows (Carter 2015), which may be particularly critical for non-standard commodity asset hedging during non-standard trading hours.

Our study has important implications for market participants, particularly for commodity merchants who use financial derivatives for risk management and inventory decisions. Swaps are effective though pricey tools to manage the risk in cash positions held during relatively long periods. At the same time, our findings are relevant for market design and surveillance. Relative to values reported in analyses conducted shortly after the launch of the SDRs, the swaps market has significantly grown. Not surprisingly, this suggests that transparency may have contributed to ensure the swaps market success as an avenue for hedging and price discovery. In addition, the evening future market trading hours can be of great importance for swap dealers to hedge their risks from swap trading. Relevant policies to promote evening session market liquidity can facilitate easier and more timely swap dealers' hedging, which will likely drive down swap transaction costs and boost swap trading.

Further research is needed to investigate in more detail the role of swap markets. SDR data do not identify the nature of market players which would provide a better understanding of the motivations underlying participation in the swaps market. The swaps market growth generated concerns that excess speculation in these markets was causing adverse spillover effects on futures markets (e.g. Masters 2008; Masters and White 2008; Sanders and Irwin 2010; Adjemian and Plato 2010). Our price discovery and hedging analyses do not raise any alarms consistent with excess speculation. However, trader identity data would allow a more careful investigation into this issue.

4.8. Tables and Figures

Table 4.1: Summary Statistics (2018/01 to 2019/12)

	Daily Swap Volume in futures equivalent number of contracts			
	Mean	Std.	Min	Max
Corn	30239	9788	7897	62891
Soybean	20790	7711	8711	38467
Corn (USDA)	30965	11035	10325	59365
Soybean (USDA)	21168	11365	10235	35095
	Swap to Nearby Futures Contract Volume Ratio			
Corn	0.168	0.044	0.068	0.390
Soybean	0.199	0.053	0.069	0.453
Corn (USDA)	0.132*	0.102	0.052	0.277
Soybean (USDA)	0.172**	0.095	0.103	0.309
	Swap Trade Size in futures equivalent number of contracts			
Corn	47	30	19	102
Soybean	54	42	25	116
Corn (USDA)	57*	42	38	99
Soybean (USDA)	61*	51	32	107
	Swap/Futures Trade Size Ratio			
Corn	5.926	1.003	2.303	17.360
Soybean	12.132	3.362	4.325	20.325
Corn (USDA)	6.223	2.035	4.035	16.362
Soybean (USDA)	13.303	4.303	5.305	18.932
	Swap Tenor			
Corn	126 days	61 days	47 days	413 days
Soybean	141 days	40 days	39 days	226 days
Corn (USDA)	168 days*	110 days	43 days	365 days
Soybean (USDA)	175 days*	109 days	98 days	237 days
	Swap Realized Volatility (Futures Realized Volatility)			
Corn	5.979 (4.627)	2.036 (1.981)	1.362 (1.250)	22.325 (19.597)
Soybean	10.303 (9.175)	5.302 (3.122)	5.035 (3.142)	21.035 (25.377)
Corn (USDA)	6.009 (7.103)***	2.309 (1.901)	1.352 (3.932)	20.362 (21.170)
Soybean (USDA)	11.352 (13.325)***	5.035 (4.032)	3.035 (5.695)	20.362 (22.325)

Table 4.1 (cont.)

Notes: This table presents summary statistics of the swap relative to the futures markets and includes all trading hours (both day and evening). The daily volume shows the average of the futures equivalent swap trading volume across all days. We convert the swap volume (in bushels) to futures contract equivalent volume (5000 bushels per contract). The swap to future volume ratio is measured each day, using futures equivalent volume divided by nearby futures volume in the same day, then averaged across all trading days in sample. The average trade size and tenor are mean values of the daily averages. Tenors are presented in calendar days. We calculate the realized volatility from 5-minute sampled returns for both markets. The last transaction price is used for each 5-minute interval when calculating returns. In swap markets, we use the swap-implied nearby price derived from the cost-of-carry framework. The non-announcement days and USDA announcement days results are presented separately. Announcement days include any of the days in which WASDE, Crop Production, Prospective plantings, Grain Stocks, and the Acreage reports are released. There are 34 announcement days during our sample period. The asterisks indicate statistical significance when testing for the null hypothesis that the means are identical between announcement days and other days using the Welch's t-test: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 4.2: Intraday(5-minute) Information Leadership Share of Corn and Soybean Market

	Futures ILS	Swap ILS	$\frac{\text{Swap Volume}}{\text{Nearby Futures Volume}}$
(Day Sessions Only)			
Corn	0.829	0.171	0.169
Corn (USDA)	0.878	0.122	0.122
Soybeans	0.817	0.183	0.191
Soybean (USDA)	0.835	0.165	0.177
(Evening Sessions Only)			
Corn	0.679	0.321	0.333
Corn (USDA)	0.714	0.286	0.293
Soybeans	0.593	0.407	0.410
Soybean (USDA)	0.737	0.263	0.304

Notes: This table reports the information shares of the nearby futures market versus the swap market. The 2 price series are resampled every 5-minutes. Reported values are daily averages from January/2018 to December/2019. There are total of 392 trading days within the sample period. We use the Putnins (2013) information leadership share (ILS) method to estimate the information shares. The last column reports the average daily volume ratios between swap and futures markets. The evening trading session during our sample period is from 19:00 to 7:45. We have an average of 58 (207) pairs of observations for each day (evening) trading session.

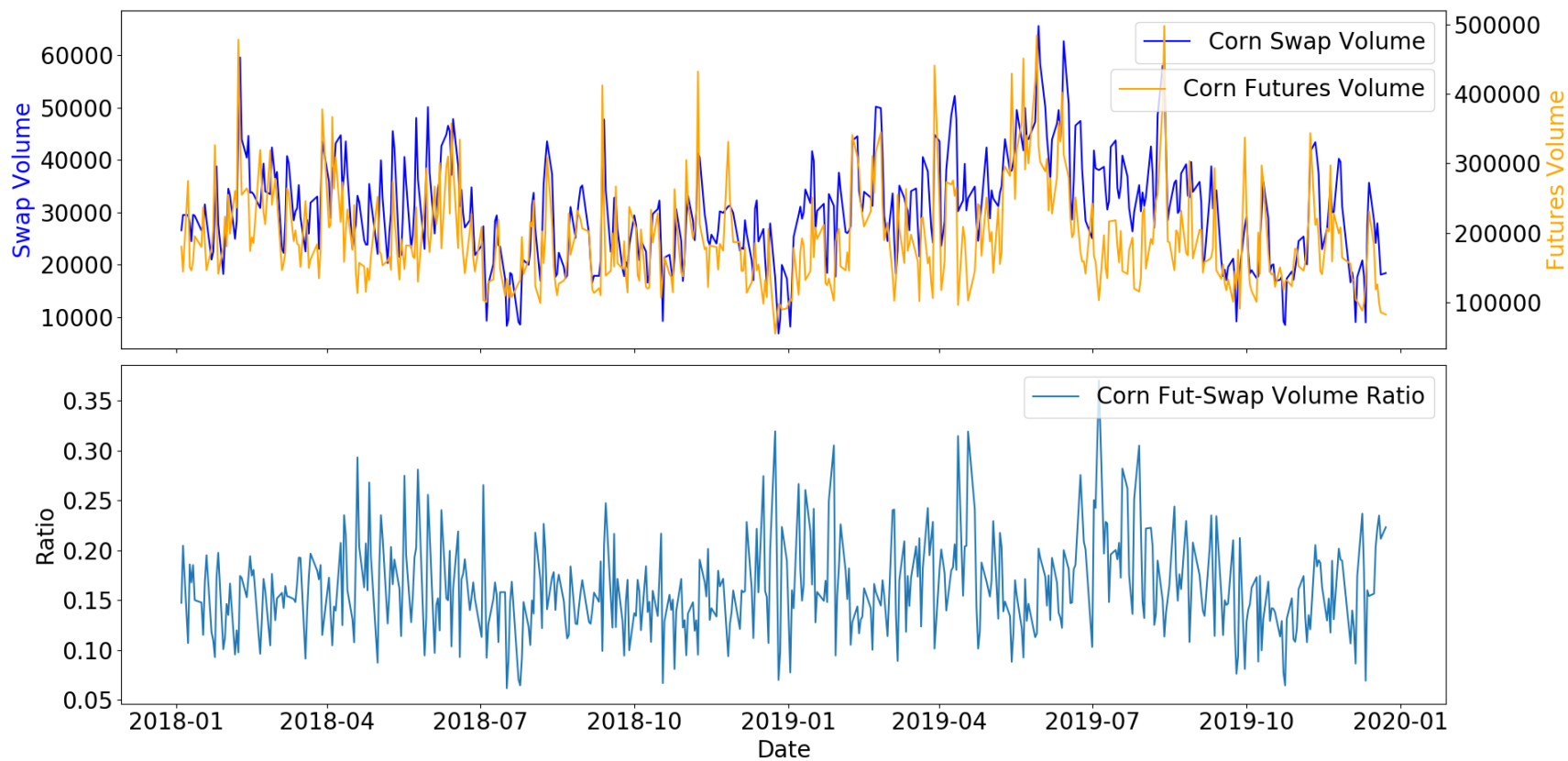


Figure 4.1: Daily Trading Volume in the Corn Swap and Futures Nearby Contract Markets

Notes: This figure shows the daily corn swap and nearby futures market volume in the top plot. The left (right) y-axis represents the swap (futures) market volume. The bottom plot shows the daily swap to nearby futures volume ratio. Correlation between the swap volume and nearby futures market contract volume is 0.505.

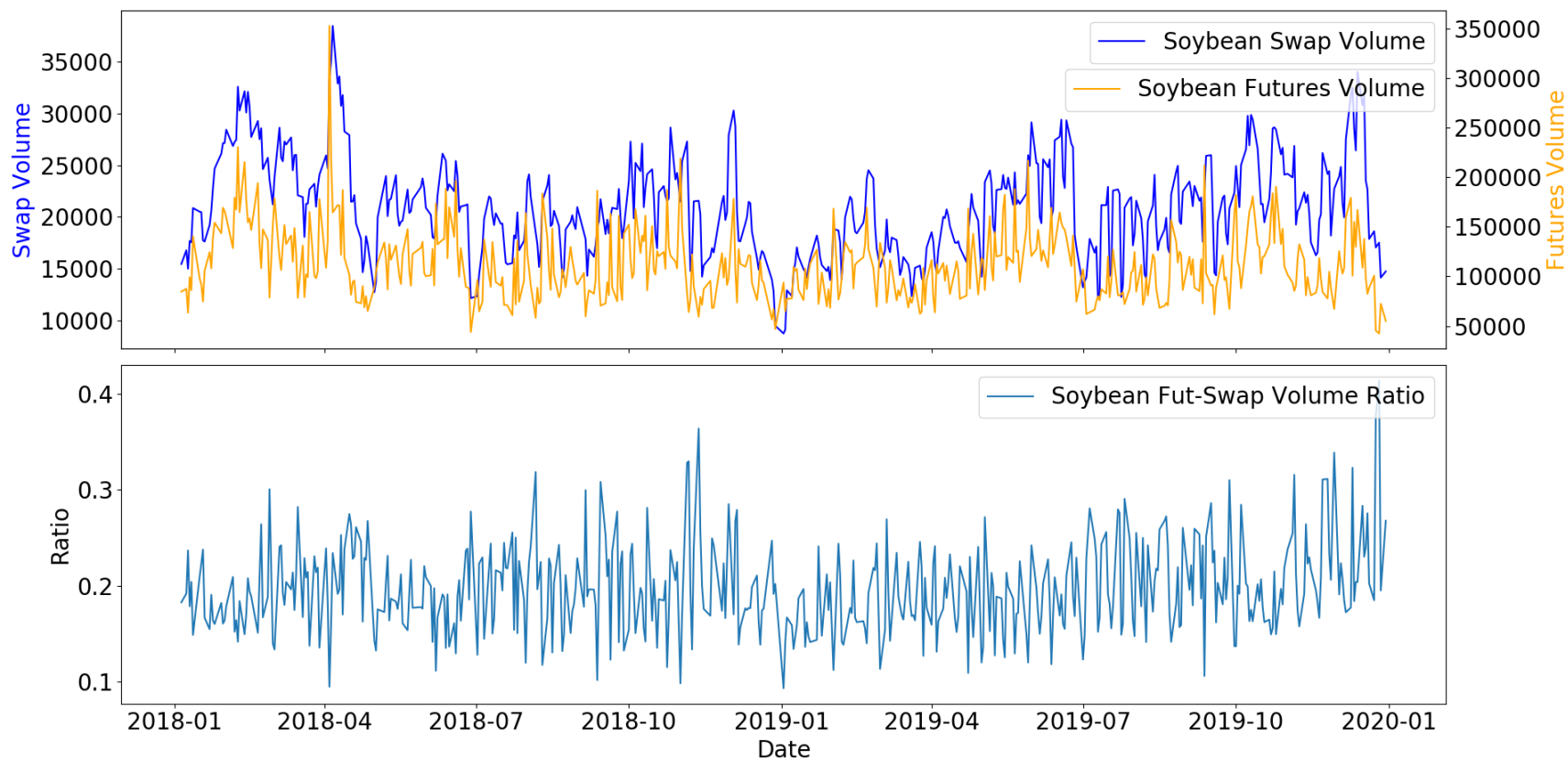


Figure 4.2: Daily Trading Volume in the Soybean Swap and Futures Nearby Contracts

Notes: This figure shows the daily soybean swap and nearby futures market volume in the top plot. The left (right) y-axis represents the swap (futures) market volume. The bottom plot shows the daily swap to nearby futures contract volume ratio. Correlation between the swap volume and nearby futures market contract volume is 0.477.

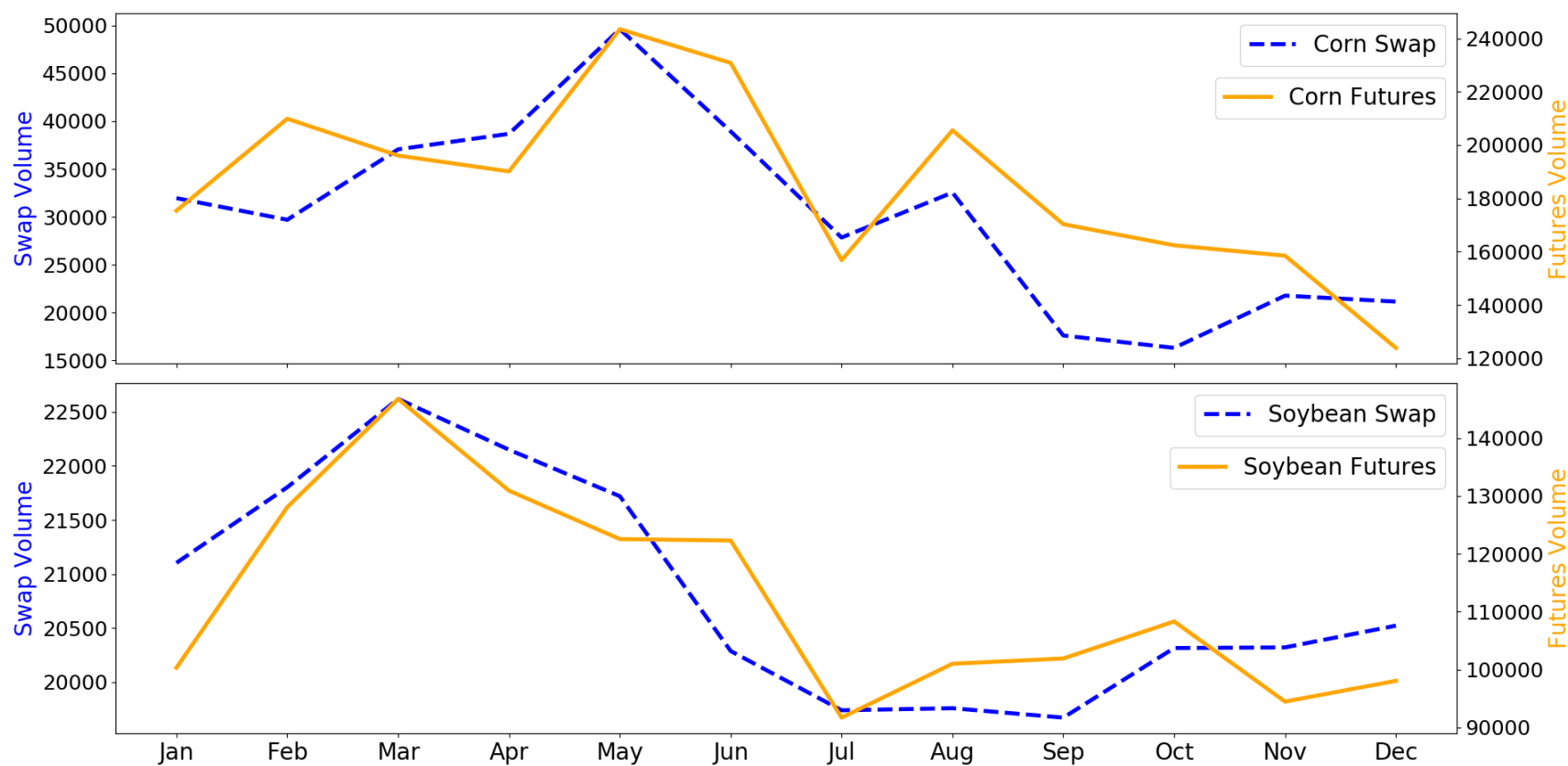


Figure 4.3: Corn and Soybean Daily Swap Volume by Month

Notes: The average daily volume is the mean of the daily volume within each month. The plot shows the average for the sample period (2018 and 2019). The left (right) y-axis represents the swap (futures) market volume.

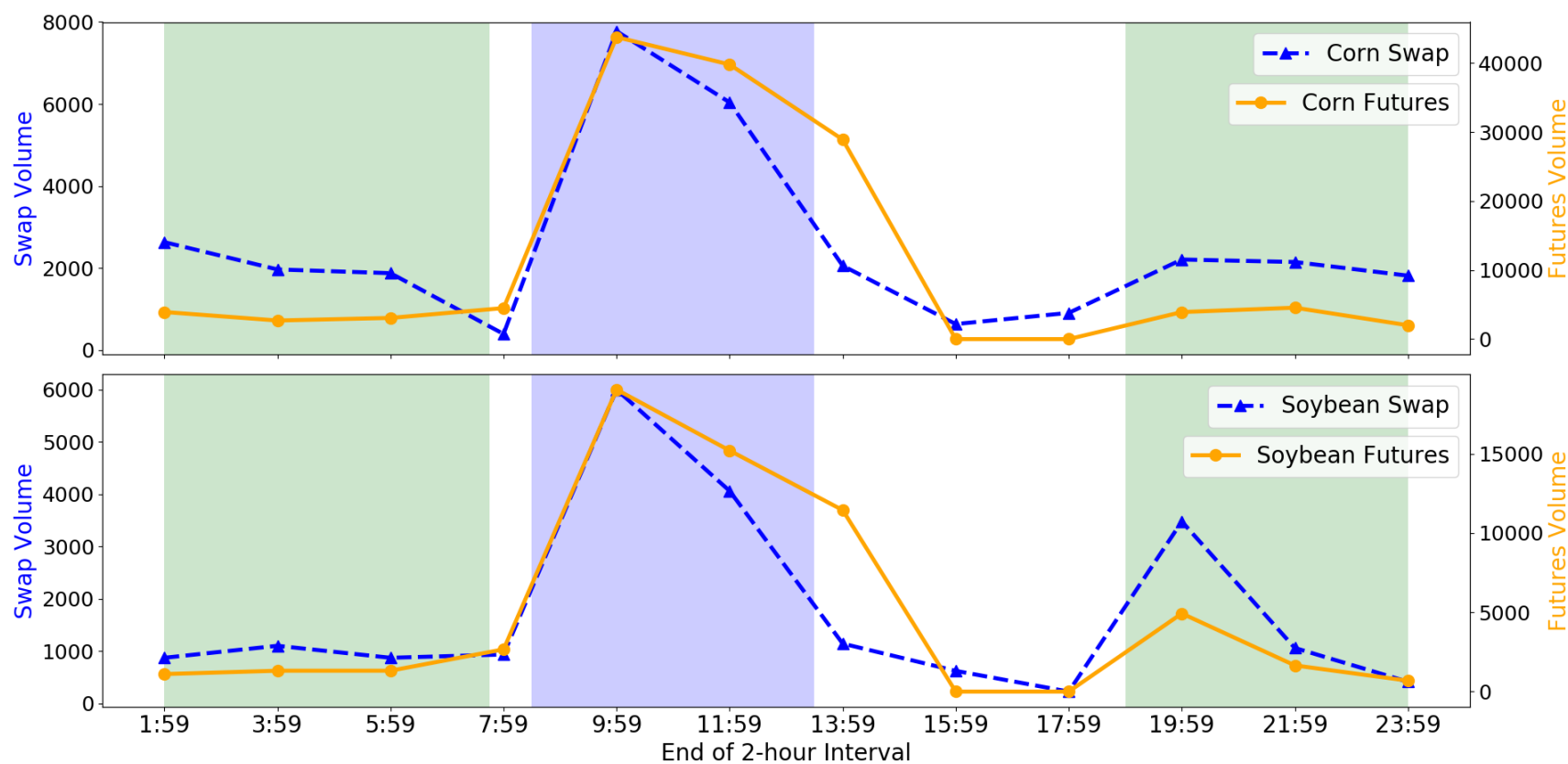


Figure 4.4: Corn and Soybeans Intraday Volume Distribution (2-hour intervals)

Notes: The plot presents the daily average of the swap market volume within each 2-hour window. The x-axis shows the end of each 2-hour window in U.S. Central Standard Time. The left (right) y-axis represents the total number of contracts within each time window for the swap (futures) market, calculated for the sample period (2018 and 2019). The blue (green) shaded area reflects the futures day (evening) trading session in both markets.

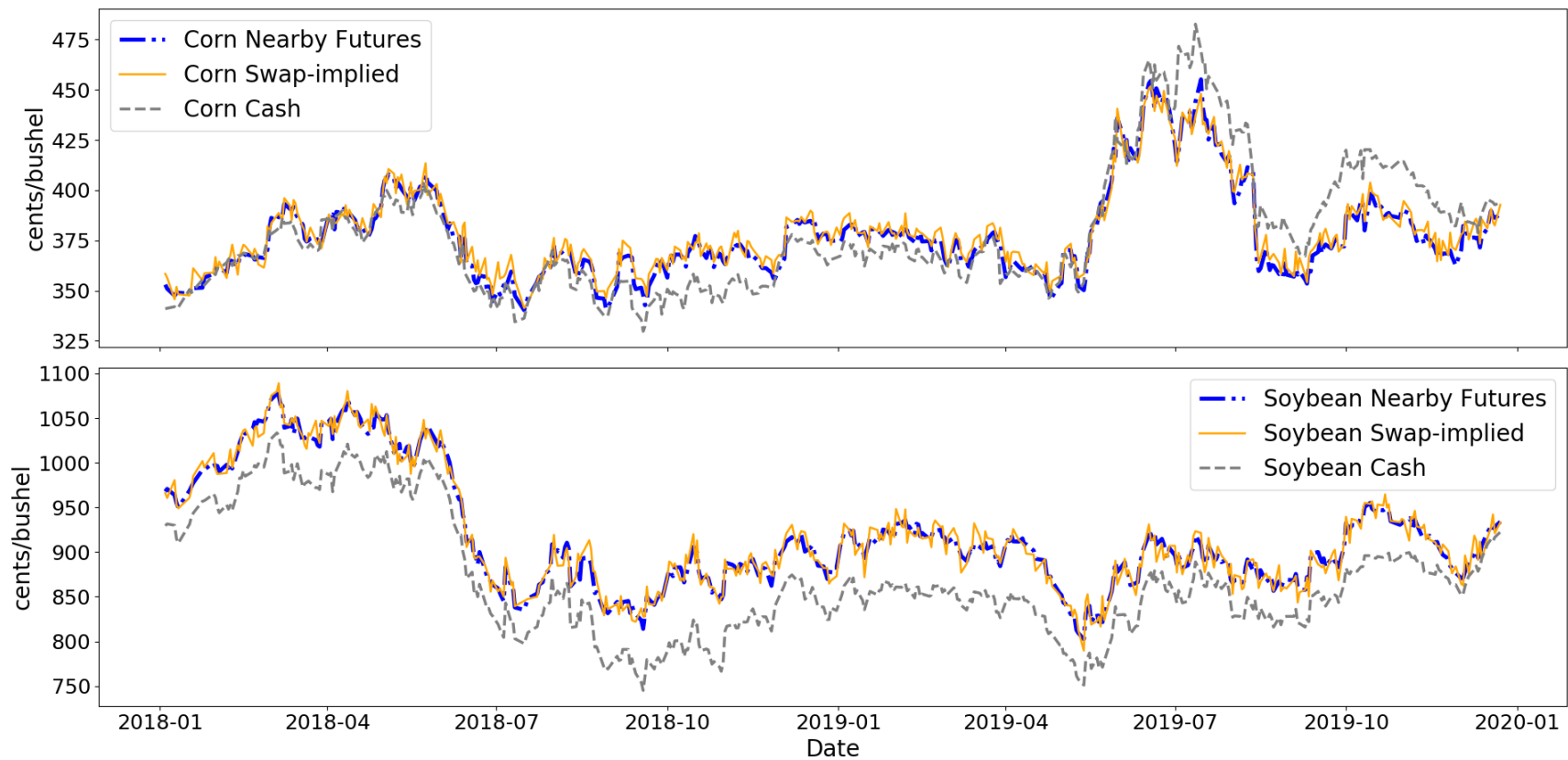


Figure 4.5: Corn and Soybean Daily Prices

Notes: This figure shows the corn (top plot) and soybean (bottom plot) market daily cash, nearby futures, and the swap implied nearby prices. The nearby futures and swap-implied prices are calculated from intraday volume-weighted prices.

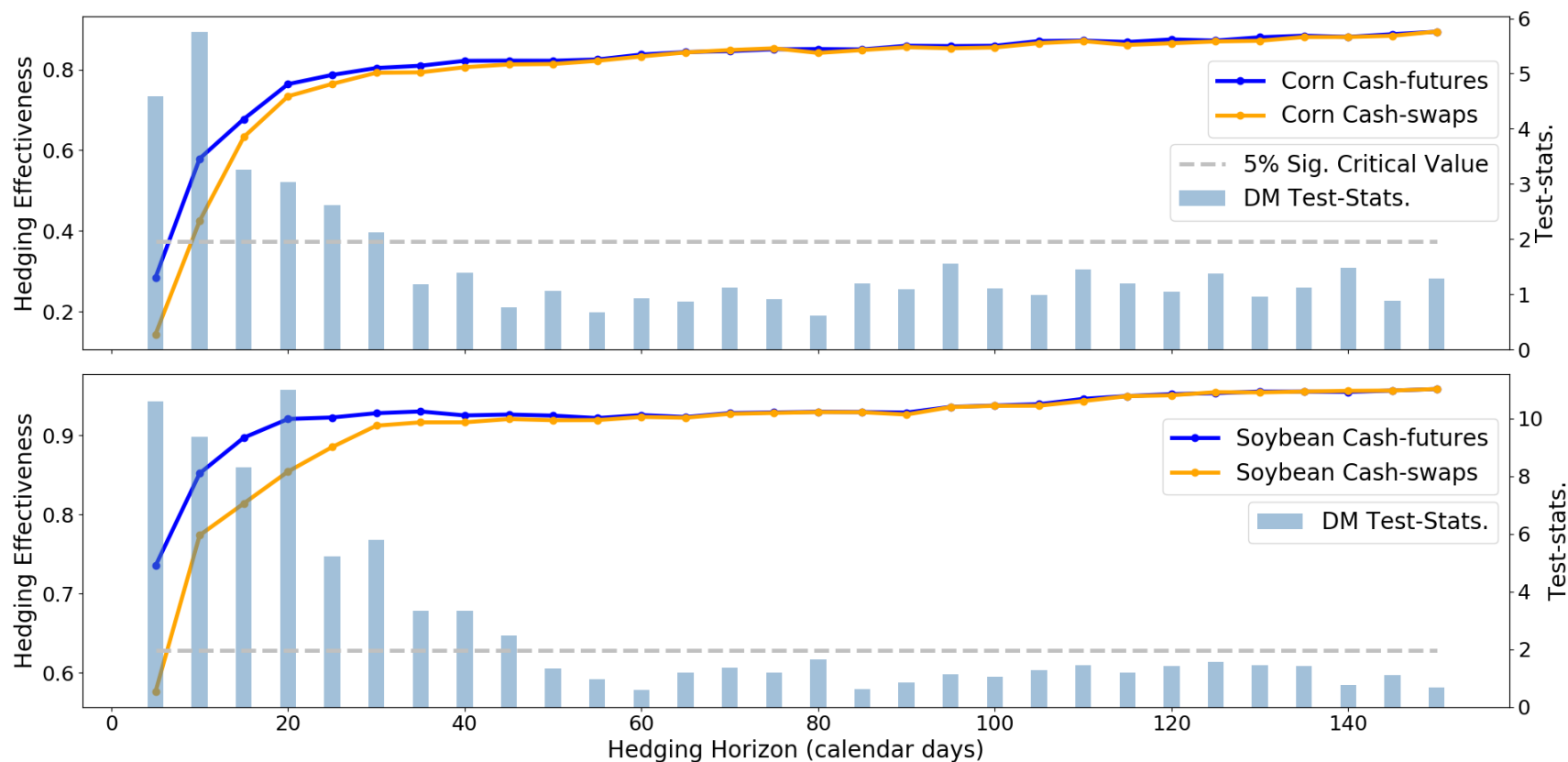


Figure 4.6: Hedging Effectiveness (HE) in Corn and Soybean Swap and Futures Markets

Notes: This plots demonstrates the hedging effectiveness (HE) of corn (top plot) and soybean (bottom plot) futures and swap contracts to protect against cash price risks from Jan-2018 to Dec-2019. The blue (orange) line represents the cash-futures (cash-swap) HE over different hedging horizons (5 to 150 days in increases of 5 days). We conduct the DM (Diebold and Mariano 1995) test to examine the difference between the two hedging strategies. The blue bars demonstrate the absolute values of the DM test statistics after controlling for non-*iid* errors (Newey and West 1987). The silver dashed line is the critical value (5% significance level) for the null hypothesis that the two hedging strategies are not statistically different.

4.9. References

- Adjemian, M.K. and G.E. Plato. 2010. The Dodd-Frank Wall Street Reform and Consumer Protection Act: Changes to the Regulation of Derivatives and Their Impact on Agribusiness. AIS-89, Economic Research Service, U.S. Department of Agriculture. Available at <http://www.ers.usda.gov/publications/ais-agricultural-income-and-finance-outlook/ais89.aspx>
- Adjemian, M.K., Smith, A. and He, W., 2021. Estimating the market effect of a trade war: The case of soybean tariffs. *Food Policy*, 105, p.102152.
- Ackerman, A. 2013. “CFTC Misreporting Size of Swaps Market, Agency Says.” *The Wall Street Journal*, December 18, 2013. Available at <http://online.wsj.com/news/articles/SB10001424052702304866904579266851056302512>
- Ackerman, A. and K. Burne. “Acknowledging Mistake, U.S. Regulators Still Struggle to Oversee Derivatives Market.” *The Wall Street Journal*, May 1, 2014. Available at <http://online.wsj.com/news/articles/SB10001424052702303948104579536251048387342>
- Aulerich, N.M., Irwin, S.H. and Garcia, P., 2014. 6. Bubbles, Food Prices, and Speculation: Evidence from the CFTC’s Daily Large Trader Data Files (pp. 211-260). University of Chicago Press.
- Arzandeh, M. and Frank, J., 2019. Price Discovery in Agricultural Futures Markets: Should We Look beyond the Best Bid-Ask Spread?. *American Journal of Agricultural Economics*, 101(5), pp.1482-1498.
- Baillie, R.T., Booth, G.G., Tse, Y. and Zabotina, T., 2002. Price discovery and common factor models. *Journal of Financial Markets*, 5(3), pp.309-321.
- Benos, E., Payne, R. and Vasios, M., 2020. Centralized trading, transparency, and interest rate

- swap market liquidity: Evidence from the implementation of the Dodd–frank act. *Journal of Financial and Quantitative Analysis*, 55(1), pp.159-192.
- Beveridge, S. and Nelson, C.R., 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), pp.151-174.
- Carter, C., 2015. *Futures and Options Markets: An Introduction*. Rebel Text.
- CME, 2021. China’s High Demand for Soybeans Fuels Asian Hours Futures Trading. Available at: <https://www.cmegroup.com/education/articles-and-reports/chinas-high-demand-for-soybeans-fuels-asian-hours-futures-trading.html>
- Commodity Futures Trading Commission (CFTC). Commodity Swap Dealers & Index Traders with Commission Recommendations. CFTC Staff Report. Washington, DC, September 2008.
- . 2011. Q&A – Final Rulemaking on Swap Data Recordkeeping and Reporting. Available at http://www.cftc.gov/ucm/groups/public/@newsroom/documents/file/sdrr_qa_final.pdf
- . 2012. Q&A – On Start of Swap Data Reporting. Available at http://www.cftc.gov/ucm/groups/public/@newsroom/documents/file/startreporting_qa_final.pdf
- . 2013. Division of Market Oversight Advisory, March 8, 2013. Available at <http://www.cftc.gov/ucm/groups/public/@newsroom/documents/file/dmo advisory030813.pdf>
- . 2022. “Swap Execution Facilities (SEFs).” Available at: <https://www.cftc.gov/IndustryOversight/TradingOrganizations/SEF2/index.htm>
- Couleau, A., Serra, T. and Garcia, P., 2020. Are corn futures prices getting “jumpy”? *American*

- Journal of Agricultural Economics*, 102(2), pp.569-588.
- Diebold, F.X., 2015. Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold–Mariano tests. *Journal of Business & Economic Statistics*, 33(1), pp.1-1.
- Diebold, F. and Mariano, R., 1995. Comparing predictive accuracy. *Journal of Business & Economics Statistics* 13:253–263.
- Ederington, L.H., 1979. The hedging performance of the new futures markets. *The Journal of Finance*, 34(1), pp.157-170.
- Engle, R.F. and Granger, C.W., 1987. Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, pp.251-276.
- Garbade, K.D. and Silber, W.L., 1983. Price movements and price discovery in futures and cash markets. *The Review of Economics and Statistics*, pp.289-297.
- Giacomini, R. and White, H., 2006. Tests of conditional predictive ability. *Econometrica*, 74(6), pp.1545-1578.
- Hansen, P.R. and Lunde, A., 2006. Realized variance and market microstructure noise. *Journal of Business & Economic Statistics*, 24(2), pp.127-161.
- Hasbrouck, J., 1995. One security, many markets: Determining the contributions to price discovery. *The Journal of Finance*, 50(4), pp.1175-1199.
- Hasbrouck, J., 2021. Price discovery in high resolution. *Journal of Financial Econometrics*, 19(3), pp.395-430.
- He, X. and Serra, T., 2022. Are price limits cooling off agricultural futures markets?. *American Journal of Agricultural Economics*. Forthcoming.
- Hu, Z., Mallory, M., Serra, T. and Garcia, P., 2020. Measuring price discovery between nearby

- and deferred contracts in storable and nonstorable commodity futures markets. *Agricultural Economics*, 51(6), pp.825-840.
- Huang, J., Serra, T., Garcia, P. and Irwin, S.H., 2022. To batch or not to batch? The release of USDA crop reports. *Agricultural Economics*, 53: 143-154.
- Hubbs, T. "Questions about Corn Crop Continue." farmdoc daily (9):153, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, August 19, 2019.
- Hull, J.C., 2015. Options futures and other derivatives. Pearson Education.
- Irwin, S., and T. Hubbs. "Late Planting and Projections of 2019 U.S. Corn and Soybean Acreage." farmdoc daily (9): 85, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, May 9, 2019.
- Irwin, S.H. and Sanders, D.R., 2010. The impact of index and swap funds on commodity futures markets: OECD Library <https://www.oecd-ilibrary.org/content/paper/5kmd40wl1t5f-en>.
- Janzen, J.P. and Adjemian, M.K., 2017. Estimating the location of world wheat price discovery. *American Journal of Agricultural Economics*, 99(5), pp.1188-1207.
- Keele, L. and De Boef, S., 2004. Not just for cointegration: error correction models with stationary data. Working paper, available at: <https://www.nuff.ox.ac.uk/Politics/papers/2005/Keele%20DeBoef%20ECM%20041213.pdf>
- Lehecka, G.V., Wang, X. and Garcia, P., 2014. Gone in ten minutes: Intraday evidence of announcement effects in the electronic corn futures market. *Applied Economic Perspectives and Policy*, 36(3), pp.504-526.
- Loon, Y., and Zhong, Z., 2016, Does Dodd-Frank affect OTC transaction costs and liquidity?

- Evidence from real-time CDS trade reports, *Journal of Financial Economics* 119, 645–672.
- Maples, W.E., Brorsen, B.W. and Etienne, X.L., 2019. Hedging effectiveness of fertilizer swaps. *Applied Economics*, 51(53), pp.5793-5801.
- Masters, M.W., 2008. Testimony before the committee on homeland security and governmental affairs. US Senate, Washington, May, 20.
- Masters, M.W. and White, A.K., 2008. The accidental Hunt brothers: How institutional investors are driving up food and energy prices. The Accidental Hunt Brothers Blog, special report posted July, 31, p.2009.
- Miedema, D. 2014 “Market Operators Hinder CFTC As It Seeks to Monitor Swaps – Sources.” Reuters, March 25, 2014. Available at <http://www.reuters.com/article/2014/03/25/usderivatives-policing-idUSBREA2O1YT20140325>
- Mixon, S., Onur, E. and Riggs, L., 2018. Integrating swaps and futures: A new direction for commodity research. *Journal of Commodity Markets*, 10, pp.3-21.
- Mixon, S. and Onur, E., 2020. Risk appetite and intermediation by swap dealers. Available at SSRN 3566003.
- Nazareth, A.L. and Rosenberg, G.D., 2013. The new regulation of swaps: A lost opportunity. *Comparative Economic Studies*, 55(3), pp.535-548.
- Newey, W.K. and West, K.D., 1987. Hypothesis testing with efficient method of moments estimation. *International Economic Review*, pp.777-787.
- Peterson., P. E. 2014. “How Large Is the Agricultural Swaps Market?” Proceedings of the NCCC-

- 134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO. [<http://www.farmdoc.illinois.edu/nccc134>].
- Peterson, P.E., 2018. *Commodity Derivatives: A Guide for Future Practitioners*. Routledge.
- Putniņš, T.J., 2013. What do price discovery metrics really measure?. *Journal of Empirical Finance*, 23, pp.68-83.
- Prager, D., Burns, C., Tulman, S. and MacDonald, J., 2020. Farm Use of Futures, Options, and Marketing Contracts, USDA Working Paper (No. 1473-2020-854).
- Riggs, L., Onur, E., Reiffen, D. and Zhu, H., 2020. Swap trading after Dodd-Frank: Evidence from index CDS. *Journal of Financial Economics*, 137(3), pp.857-886.
- Sanders, D.R., Irwin, S.H. and Merrin, R.P., 2010. The adequacy of speculation in agricultural futures markets: too much of a good thing?. *Applied Economic Perspectives and Policy*, 32(1), pp.77-94.
- Smith, A., 2005. Partially overlapping time series: A new model for volatility dynamics in commodity futures. *J. Appl. Economet.* 20(3), 405–422.
- Stock, J.H. and Watson, M.W., 1988. Testing for common trends. *Journal of the American statistical Association*, 83(404), pp.1097-1107.
- Stoll, H.R. and Whaley, R.E., 2010. Commodity index investing and commodity futures prices. *Journal of Applied Finance*, 20(1).
- Swanson, K., J. Coppess, G. Schnitkey and C. Zulauf. "The Trade Conflict – Impact on Illinois Agriculture in 2018." *farmdoc daily* (9):75, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, April 25, 2019.
- Wang, X., Garcia, P., & Irwin, S. H. 2014. The behavior of bid-ask spreads in the electronically-traded corn futures market. *American Journal of Agricultural Economics*, 96(2), 557-577.

- Wang, Y., Wu, C. and Yang, L., 2015. Hedging with futures: Does anything beat the naïve hedging strategy?. *Management Science*, 61(12), pp.2870-2889.
- Working, H., 1961. New concepts concerning futures markets and prices. *The American Economic Review*, 51(2), pp.160-163.
- Yan, B. and Zivot, E., 2007, February. The dynamics of price discovery. In AFA 2005 Philadelphia Meetings.
- Yan, B. and Zivot, E., 2010. A structural analysis of price discovery measures. *Journal of Financial Markets*, 13(1), pp.1-19.

4.10. Supplementary Results

4.10.1. Swap Trade Size and Tenor Distributions

Table 4.3: Percentile Distribution of Swap Trade Size and Tenors (all trades considered)

	Swap Size		Swap Tenor Length (Calendar Days)	
	Corn	Soybean	Corn	Soybean
10%	5	11	29	31
30%	27	32	105	122
50%	41	47	192	205
70%	89	92	292	309
90%	192	217	405	522
Mean	47	55	126	141

Notes: This table shows the swap trade size and tenors corresponding to different percentiles. The swap size is expressed as the futures-equivalent contract sizes, rounded to the nearest integer.

4.10.2. Swap Market Information Leadership Shares of Corn and Soybean (Timestamp Mismatch Test)

Table 4.4 Swap Market Information Shares (Mismatched Timestamp)

	1-min Mismatch	3-min Mismatch	5-min Mismatch
Day Sessions Only			
Corn	0.187	0.199	0.234
Soybeans	0.201	0.220	0.256
Evening Sessions Only			
Corn	0.339	0.356	0.387
Soybeans	0.416	0.428	0.443

Notes: This table reports the swap market information shares. The futures and swap price series are resampled every 5-minutes. Reported values are daily averages from January/2018 to December/2019. There are total of 392 trading days within the sample period. We use the Putnins (2013) information leadership share (ILS) method to estimate the information shares. To control for potential swap trade reporting delays, we reprocess (only) the swap trade timestamps by moving each timestamp 1-, 3-, or 5-minute earlier than its original time (i.e. in the 1-minute case, 8:30:00 becomes 8:29:00). We match the futures and swap market trades with re-processed timestamps and estimate the information shares. The second, third, and fourth column in this table report the swap market price discovery shares with 1-, 3-, and 5-minute mismatch scenario, respectively.

4.10.3. Price Discovery Model Explanation

The ILS measure is derived from the IS and CS models. Consider a single asset trading in n distinct markets. Let $P_t = (P_{1,t}, P_{2,t}, \dots, P_{n,t})'$ denote the vector of n prices from these different markets at time t . We assume these prices are cointegrated with cointegration rank $r = n - 1$ which involves prices have a common stochastic trend with a permanent (f_t) and a transient (G_t) component given by $P_t = f_t + G_t$ (Stock and Watson 1988). Using matrix notation, the VECM model for the n prices can be written as:

$$\Delta P_t = \alpha(\beta' P_{t-1} - \mu) + \sum_{k=1}^K A_k \Delta P_{t-k} + \epsilon_t, \quad (4.5)$$

where Δ represents the first order difference, K is the number of lags in the model and A_k is a matrix of autoregressive coefficients which measure the short-run price dynamics. ϵ_t is a vector of white noise error terms with variance-covariance matrix Ω . The $n \times r$ matrix α measures the speed of adjustment when there exist deviations from long-run equilibrium relationships. β represents the $n \times r$ matrix of coefficients of the cointegration process and is assumed to have the following form $[1: -I]$, where $-I$ is a $(n-1)$ identity matrix and 1 is a $(n-1)$ vector of 1's. Following Beveridge and Nelson (1981), the VECM model in (4.5) can be expressed as

$$P_t = P_0 + \Psi(1) \sum_{j=1}^t \epsilon_j + \Psi^*(L) \epsilon_t, \quad (4.6)$$

where matrices $\Psi(1)$ and $\Psi^*(L)$ are $n \times n$ matrices in the lag operator L , with $\Psi(1)\epsilon_t$ measuring the cumulative (long-run) impacts of an innovation ϵ_t on all future price movements. Hasbrouck (1995) and Engle and Granger (1987) demonstrate that $\beta' \Psi(1) = 0$ and $\Psi(1)\alpha = 0$. Also, the assumption of $r = n - 1$ results in $\Psi(1)$ having a rank of one. As a result, an innovation ϵ_t has the same long-run impacts on each price. $\Psi^*(L)$ measures price changes resulting from transient shocks. The variance of the common factor innovations is given by $var(\psi \epsilon_t) = \psi \Omega \psi'$, where ψ is a common row vector in the $\Psi(1)$ matrix from equation (4.6). The IS and CS models

decompose the common factor variance into the contributions of variance from different prices during the formation of the efficient price. The CS measure of the i 's price can be calculated as

$$CS_i = \frac{\psi_i}{\sum_{i=1}^n \psi_i} , \quad (4.7)$$

and the IS of the i 's price can be calculated as

$$IS_i = \frac{[\psi_i \sigma_i]^2}{\psi \Omega \psi'} , \quad (4.8)$$

where σ_i is the standard deviation in the variance-covariance matrix of price i . Hasbrouck (1995) suggests that the variance-covariance matrix is often non-diagonal, and the price innovations are strongly correlated with each other. Therefore, a Cholesky factorization is applied to the covariance matrix to eliminate contemporaneous correlation. With the Cholesky factorization, the information shares computed from (4.8) depend on the ordering of the three prices. Following previous studies, we compute the average IS by taking all possible different orderings from all prices (Hasbrouck 1995; Baillie et al. 2002; Hu et al. 2020). After the Cholesky factorization, $\Omega = \mathbf{M}\mathbf{M}'$, where $\mathbf{M}$ is a lower triangular matrix, the IS measure can be expressed as:

$$IS_i = \frac{[\psi \mathbf{M}]_i^2}{\psi \Omega \psi'} . \quad (4.9)$$

The IS and CS models decompose the common factor variance into contributions from different prices during the formation of the efficient price. This allows separating price changes into a permanent component associated with the fundamental value which follows a random walk, and a transitory component related to different speed of price adjustments. Based on the IS and CS, Yan and Zivot (2010) create the information leadership (IL) measure that nets out the transitory frictions which cause noise:

$$\begin{aligned} IL_{fut} &= \left| \frac{IS_{fut}}{IS_{swap}} \frac{CS_{swap}}{CS_{fut}} \right| \\ IL_{swap} &= \left| \frac{IS_{swap}}{IS_{fut}} \frac{CS_{fut}}{CS_{swap}} \right| \end{aligned} \quad (4.10)$$

The IL_{fut} and IL_{swap} measure the information leadership for futures and swap markets, respectively. However, the IL metrics from Yan and Zivot (2010) do not add up to 1 and thus cannot be interpreted as a share. For easier comparability and interpretation, Putnins (2013) develops an information leadership share (ILS) based on the following normalized IL measures:

$$ILS_{fut} = \frac{IL_{fut}}{IL_{fut} + IL_{swap}}, \quad ILS_{swap} = \frac{IL_{swap}}{IL_{fut} + IL_{swap}}. \quad (4.11)$$

Both ILS_{fut} and ILS_{swap} move in the range $[0,1]$, and $ILS_{fut} + ILS_{swap} = 1$.

CHAPTER 5

CONCLUSIONS

This dissertation is timely and identifies three circumstances in which market quality can be compromised that have not been previously investigated. These include futures' market design allowing for a certain degree of opacity, existence of persistent arbitrage opportunities in the futures markets, and the role of the still rather opaque swaps market in forward trading and its interactions with futures markets. This dissertation has three specific objectives: (1) to examine the impact hidden limit orders (HLOs), a form of market opacity, on the futures' market quality; (2) to test whether persistent mispricing and trending exist in futures markets; and (3) to examine the relevance, price discovery, and hedging effectiveness of swap markets relative to the futures. The dissertation consists of three stand-alone essays that focus on each of these objectives.

In the first essay, we investigate for the first time the presence of hidden liquidity (HL) in agricultural futures markets and its consequences for market quality. Our conservative estimates indicate that HLOs represent more than 10% (20%) of the total trading volume in corn (live cattle) futures market. Our findings suggest that the existence of HLOs improves market quality in multiple dimensions: increasing trading volume (7% in corn and 5% in live cattle) while reducing market volatility by at least 5% and reducing the relative trading spread by at least 11%. While the classic microstructure theory does not support opacity as it can lead to diminished quality, our research does not support this view. This essay suggests that HLOs can successfully serve the purpose of slow traders (i.e. farmers, grain commercials) and improve market quality. Additionally, hidden liquidity can provide a signal for slower traders of subsequent market quality conditions, allowing them to establish market positions.

The second essay investigates the futures market calendar spread time series momentum (STSM) in 22 U.S. commodity futures markets during the period from 1986 to 2019. For the first time in the momentum literature, we adopt a comprehensive approach that corrects for both empirical design and econometric issues which were pervasive in previous research. Since STSM profitability is predicated on return predictability, we use regression analysis to shed light on this issue. We examine whether historical spread returns can predict future spread and outright returns. Results demonstrate that return predictability is weak at the market-level. Although evidence of return predictability can be found with the pooled regression, we show that it is mainly driven by inclusion of unrealizable roll yields, with statistical issues playing a lesser role. The latter are induced by heterogeneity in return mean and volatility across different markets. STSM does not generate returns that are statistically different from the benchmark long-only trading strategy, with both strategies generating very low or negative returns. For the momentum to outperform the benchmark strategy, predictability should be at least three times larger than observed from real data but would entail substantial downside risk. In sum, the empirical evidence indicates that returns from STSM-type strategies are illusive for the commodities and period studied, thus suggesting that markets have continued to operate efficiently.

The third essay explores the role of swap contracts in commodity markets. With the real-time released swap trade data from the Swap Data Repositories (SDR), we investigate the characteristics corn and soybean swap markets relative to futures markets. In addition, we estimate the price discovery share and hedging effectiveness of swap contracts which are the two fundamental roles of financial instruments like the futures and swaps (Working 1962; Garbade and Silber 1982). Empirical evidence suggests that the overall swap market volume is 17% - 20% of the nearby futures contract volume. The price discovery shares in corn and soybean swap markets

are commensurate with their volume shares relative to the futures markets. However, compared with day trading hours, we find both the volume and price discovery share increase dramatically during evening trading hours, and this pattern is particularly strong in the soybean swap market. Further, hedging in swap contracts is as effective as hedging in futures contracts for hedging horizons longer than 35 (50) days in the corn (soybean) market. While being useful tools in price discovery and hedging, our results suggest that swap contracts have larger transactions costs than futures contracts (3.8 to 5.1-fold). Largest transactions costs are registered in the soybean market, and thus need to be understood in the framework of the U.S-China trade war and its large impacts on the soybean market.

Many issues remain to be answered in the space of commodity market quality research. First, we know little about the responses of different traders to HLOs. While some traders may speculate for profit when HLOs appear, others may take the opportunity to build or unfold their positions. Second, further research can be made to complement the second essay. A thorough examination of time-varying predictability may bring a deeper understanding of price behavior and momentum trading, especially in the context of fast electronic markets, where price inefficiencies may be arbitrated away much faster than before. Additionally, we know little about how different types of traders contribute to swap market price discovery and their potential impacts. It is also unclear whether swap markets behave any differently before and after the Dodd Frank Act. Research on these questions will permit a clearer understanding of commodity markets' quality.

5.1. References

Degryse, H., Karagiannis, N., Tombeur, G. & Wuyts, G. (2018). Two shades of opacity: Hidden orders versus dark trading. <http://dx.doi.org/10.2139/ssrn.2669447>

Frey, S., & Sandås, P. (2017). The impact of iceberg orders in limit order books. *Quarterly Journal of Finance*, 7(03), 1750007.

Garbade, K.D. and Silber, W.L., 1983. Price movements and price discovery in futures and cash markets. *The Review of Economics and Statistics*, pp.289-297.

Working, H., 1962. New concepts concerning futures markets and prices. *The American Economic Review*, 51(2), pp.160-163.