5-2022

Decision Making Under Uncertainty Among Agricultural Commodity Traders

Marei Undine Houpert
University of Arkansas, Fayetteville

Decision Making Under Uncertainty Among Agricultural Commodity Traders

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Public Policy

by

Marei Undine Houpert
University of Applied Sciences Bingen
Diplom-Betriebswirtin (FH), 2011
Ghent University
International Master of Science in Rural Development, 2013
University of Arkansas
Master of Science in Agricultural Economics, 2013

May 2022
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

Andrew McKenzie, Ph.D.
Dissertation Director

Brink Kerr III, Ph.D
Committee Member

Eric Wailes, Ph.D
Committee Member

Abstract

Understanding how humans behave has received increased attention in research in the most recent decades. While insights into human behavior are essential for the functioning of markets, these insights are also crucial for the guidance of policymaking and public spending decisions to ensure that markets are fair and accessible for all. This research has the unique opportunity to use a dataset of commercial grain trader transactions to study the decision-making and behavior of commercial traders in agricultural commodity markets rather than institutional investors or speculators.

Paper one focuses on the performance of commercial grain traders and their ability to perform above expectations consistently and the role of gender and experience in achieving their performance. First, a Fisher Exact ranking test is consulted in assessing the ability of traders to perform persistently in two consecutive periods. Next, a top and bottom rank test is used to evaluate the magnitude of their profitability over successive periods. In addition to the two tests for persistent trader performance, the influence of gender and experience were also investigated in paper one.

The second paper focuses on the use of reference points in the decision-making of commercial grain traders. For this study, a two-way panel data model was chosen to test the influence of a trader’s past performance or commodity price level, grain basis, influenced the purchasing decisions of commercial grain traders when purchasing grain with forward contracts. Past trader performance is measured in trader margin, calculated after a transaction is completed. The amount of bushels purchased and the forward contract length were investigated to capture risk in commodity markets by the size of the transaction or duration of the forward contract.

Finally, in the third paper, the role of the United States Department of Agriculture World Agriculture Supply and Demand Estimate (WASDE) report is investigated regarding its public value compared to private estimates that try to forecast the stocks-use-ratio of commodities. In this paper, we also analyze the reaction of corn producers and not only the commercial grain traders. Further, the different grain marketing seasons and weeks around the release of the WASDE were tested for their importance in making buying or selling decisions in corn markets.

Acknowledgments

I would like to express my most profound appreciation to my advisor, Dr. Andrew McKenzie, for his invaluable guidance and continuous support. He patiently helped me overcome the many difficulties during my studies ever since coming to the United States. I would like to extend my sincere thanks to Dr. Eric Wailes for believing in me, for his constructive criticism, and witty humor. Completing this dissertation would not have been possible without the support and nurturing of Dr. Brink Kerr, who has helped me in countless ways. I'd also like to acknowledge Dr. Rodolfo Nayga's insightful suggestions gratefully. Further, I am also grateful for the support of the various staff members in the Public Policy and Agricultural Economics and Business departments.

I cannot begin to express my thanks to my coworkers at Mullins Library, who are among the kindest humans I have ever met. I must especially thank my supervisor Necia Parker-Gibson for her profound belief in my work and unparalleled support during this research journey. I also had the great pleasure of working with Norma Johnson, who has been a great source of encouragement.

Never would I have thought of continuing my education without the encouragement of my dear friend Stephane Makombe. I am very grateful for the shared conversations and help in all aspects of life. I am incredibly thankful to parents Jutta and Walter Houpert for their loving support and patience. I would also like to thank all the other family members who have supported me in many ways. It is not always easy to live in a different country, but I am glad to feel loved every time I come home.

Finally, I want to thank my partner and the love of my life, Thom, who has been a constant source of encouragement and support. I am looking forward to spending my now freed-up time with him.

Contents

Introduction

Markets, where goods or services are traded, play an essential part in capitalist societies. Neoclassical economics and expected utility theory assume the self-interest of market participants as the motivation to trade and rational behavior (Von Neumann & Morgenstern, 1944). Under this assumption, a functional market can only be achieved when individual market actors feel confident enough to participate in the market and accept exposure to market risk. In an ideal case, markets provide the platform for exchange until an efficient allocation of resources is achieved, and every participant's utility is maximized before making anyone else worse off, as assumed by the Pareto optimum (Beckert, 2009). During the 2nd half of the 20th century, the efficient market hypothesis was the status quo in the finance world. The efficient market theory was characterized by Eugene Fama (1970), who stated that a market is "efficient" when prices at all times "fully reflect" available information (Fama, 2021, p.383). This theory implies that no market participant can continuously beat the market. Towards the end of the 20th century, the efficient market theory was further challenged by questioning the strength of its fundamental forces, such as arbitrage (Shleifer, 2000). Following these discoveries, behavioral finance and behavioral economics gained popularity by no longer assuming rational behavior of individuals or total efficiency of markets.

One of the most critical behavioral explanations of decision-making under uncertainty is Kahneman and Tversky's so-called prospect theory. Kahneman and Tversky wrote a scathing criticism of expected utility theory in their 1979 paper, *Prospect Theory: An Analysis of Decision Making Under Risk* (Kahneman & Tversky, 1979). They described decision-making under risk as a choice between prospects and gambles. One of their most groundbreaking advances was discovering that individuals behave differently regarding risky choices when facing losses rather than gains. Where risk aversion turns into risk-seeking behavior, "you just like winning and dislike losing – and you almost certainly dislike losing more than you like winning" (Kahneman, 2011, p.281). A second important characteristic of prospect theory is the use of reference points. Reference points play a crucial role in how humans evaluate gains and losses. For example, in the case of stocks, the current stock price, the purchase price, or the all-time high price could potentially be used to evaluate the current stock price (Nofsinger, 2017).

Prospect theory has been applied in many different fields of study today. Most prominent is the use of prospect theory in economics and business; it accounted for about half of the citations in 2000 of Kahneman and Tversky's 1979 Prospect Theory (Mercer, 2005). In economic research, prospect theory is especially suitable for studies of insurance and finance, where risk plays a central role (Barberis, 2013). Nevertheless, prospect theory has since been applied more broadly in other fields like crisis management (Feng & He, 2018) or public policy (Bellé, Cantarelli, & Belardinelli, 2018). Examples of the use of prospect theory for agricultural policy include studies of crop insurance (Babcock, 2015) or organic farming (Wang, Zhu, Zhang, & Wang, 2018).

Uncertainty in agricultural markets is a risk factor driving agricultural commodity prices that can ultimately lead to food crises. Volatility in agricultural prices is unavoidable based on the nature of agricultural production being seasonal and weather-dependent. However, it can lead to problems in food security and panic-driven market behavior. Still, the production and marketing of agricultural commodities have a long-rooted tradition in the resource-rich United States. As early as 1911, Siebel Harris pointed out that the agricultural marketing efforts each year burden the county's financial resources" (Harris, 1911, p.354). The agricultural programs continue to be costly to the United States government, with a 4.6 billion dollar budget for the commodity programs of the U.S. Department of Agriculture (USDA) and another 3.3 billion dollars for research, education, and economics (*United states department of agriculture FY 2021 budget summary.*). Concerning the massive government spending on agriculture, it is crucial to understand the behavior of the different market participants, such as producers, traders, speculators, hedgers, and investors. The functioning of commodity markets depends on the market participant behavior as it can potentially impact the functioning of the market (Coval & Shumway, 2005; Nofsinger, 2018). To better guide and educate market participants, it is crucial to understand their conscious and subconscious behavior.

Two major parties in the agricultural commodities markets are the producers, selling their products, and commercial grain traders that either function as buyers for industrial end-users or act as middlemen between the producers and industry. Both the producers and traders can use the futures market to hedge price risk. The farmers or producers can use the futures market to lock in fixed prices to sell their harvest at a later date, and vice versa, the processing industry can secure a fixed price for their future input needs. When hedging grain trades, the so-called basis represents the difference between the spot price and the price of the respective futures contract. It includes the financial cost of storage, the cost of cleaning grain to the standard of the futures contract, the cost of physical storage, including insurance, from the harvest to the actual delivery of the grain. Lastly, the basis also includes the delivery cost to or from the futures contract's delivery spot (Geman, 2005). In the grain merchandising business, the basis is more important than the price per bushel, as the basis serves as the standard method of grain valuation. The advantage of the use of basis over the bushel price is that the basis is more predictable, moving in more distinct patterns (Lorton & White, 2002).

Information plays a vital role in agricultural markets, especially price information of agricultural commodities, which influences decisions about storage, but also about the production and consumption of those commodities (Geman, 2005). (Grossman (1977) argues that the futures market serves an informational role where informed people who collect and analyze information share their future price information with uninformed traders. As a result, informed traders could earn a return on investment from gathering and processing information. The abilities to forecast spot and futures prices serve large social utility (Geman, 2005). Perloff & Rausser (1983) describe how Grossman's finding show that information has a public element, where commodity prices may be sufficient for uninformed market participants stemming from underinvestment in information. While the traditional competitive market theory assumes that all information is freely and equally available to all market participants, this assumption has been refuted for agricultural commodity markets. Ultimately asymmetric information in agricultural commodity markets can lead to increased market power (Perloff & Rausser, 1983).

The gathering, processing, and dissemination of market information serve a significant economic function. The completeness and accuracy of information make markets function more efficiently. The lack of timely fundamental market information is a known driver of agricultural price volatility that can significantly affect food security (IMF, OECD, and WFP UNCTAD, 2011). Given the role of information, the federal government, in the form of the U.S. Department of Agriculture (USDA), provides market information by publishing several periodical reports (Hieronymus, 1977). One of the prominent publications of the USDA is the World Agricultural Supply and Demand Estimates (WASDE). The WASDE provides public information in the domestic and international situation and outlook information for multiple crops (Isengildina-Massa, Irwin, Good, & Gomez, 2008). Since agricultural commodity market-related information is valuable, several private news services collect, analyze, and disseminate information to paying subscribers in an attempt to provide comparable and improved information than the WASDE does, but faster (Hieronymus, 1977; Isengildina-Massa, Karali, & Irwin, 2020). Given the increase of availability of private information and cuts in fiscal spending, the role and importance of the WASDE has been questioned since the 1980s (Just, 1983; Xie, Isengildina-Massa, Dwyer, & Sharp, 2016).

This dissertation contains three studies that seek to enhance the understanding of commercial grain trader behavior. The first study uses a winner and loser rank test and a top and bottom performance test. Similar tests were used by Aulerich, Irwin, & Garcia (2013) to gain insight into noncommercial traders' performance. The winner and loser rank tests are used to analyze the consistency of commercial trader performance. In contrast, the top and bottom test shows differences in the magnitude of the top and bottom trader performances. In addition, gender and experience differences of traders are used to analyze professional trader performance. For the second study, a two-way fixed effect panel data analysis (Jaba, Robu, & Balan, 2017) was conducted to understand how a commercial trader's past performance influences their future forward trading performance. The trading performance is measured by the amount of corn purchased and the length of forward contracts. Lastly, the third study tests the influence of the WASDE report on commercial grain traders' purchasing behavior and the selling behavior of corn producers. For this study, a time series model was selected (Brocklebank & Dickey, 2003).

Our three papers provide essential insights for commodity market research and policymakers. The dataset used for all three studies is unique as it was provided by a private company and offers insights into grain merchants' behavior. This type of data is usually not readily available.

The first paper adds to the understanding of the behavior of commercial grain traders and their performance. The study also offers insights into the repeated performance of commercial grain traders and their performance based on gender or experience. The second paper studies the past performance of commercial grain traders as an indicator of their future performance. The final paper considers the WASDE report as a reference point for commercial grain traders and corn producers. Since the WASDE is freely available online, it seems like a natural reference point to anchor grain purchasing or selling decisions on the supply and demand estimates.

Our results show that trader behavior is influenced by past performance and that reference points offered as public information may efficiently serve agricultural producers as a reference point in their decision-making. Further, our results demonstrate that the behavior of market participants is not fully understood yet. Thus, further research can help guide policymakers to effectively allocate financial resources for education about agricultural markets and distribute helpful information for market participants to make agricultural markets fair and efficient.

References

Aulerich, N. M., Irwin, S. H., & Garcia, P. (2013). Returns to individual traders in agricultural futures markets: Skill or luck? *Applied Economics*, *45*(25), 3650-3666. doi:10.1080/00036846.2012.727979

Babcock, B. A. (2015). Using cumulative prospect theory to explain anomalous crop insurance coverage choice. *American Journal of Agricultural Economics*, *97*(5), 1371-1384.

Barberis, N. C. (2013). Thirty years of prospect theory in economics: A review and assessment. *Journal of Economic Perspectives*, *27*(1), 173-196.

Beckert, J. (2009). The social order of markets. *Theory and Society, 38*(3), 245-269.

Bellé, N., Cantarelli, P., & Belardinelli, P. (2018). Prospect theory goes public: Experimental evidence on cognitive biases in public policy and management decisions. *Public Administration Review, 78*(6), 828-840.

Brocklebank, J. C., & Dickey, D. A. (2003). *SAS for forecasting time series* John Wiley & Sons.

Coval, J. D., & Shumway, T. (2005). Do behavioral biases affect prices? *The Journal of Finance, 60*(1), 1-34.

Fama, E. F. (2021). Efficient capital markets a review of theory and empirical work. *The Fama Portfolio,* 76-121.

Feng, H., & He, K. (2018). Prospect theory, operational code analysis, and risk-taking behaviour: A new model of china's crisis behaviour. *Contemporary Politics*, *24*(2), 173-190.

Geman, H. (2005). *Commodities and commodity derivates* John Whiley & Sons Ltd.

Grossman, S. J. (1977). The existence of futures markets, noisy rational expectations and informational externalities. *The Review of Economic Studies, 44*(3), 431-449.

Harris, S. (1911). Methods of marketing the grain crop. *The Annals of the American Academy of Political and Social Science, 38*(2), 36-57.

Hieronymus, T. A. (1977). Economics of futures trading for commercial and personal profit.

IMF, OECD, and WFP UNCTAD (2011). *Price volatility in food and agricultural markets: Policy responses*. FAO: Roma, Italy.

Isengildina-Massa, O., Irwin, S. H., Good, D. L., & Gomez, J. K. (2008). The impact of situation and outlook information in corn and soybean futures markets: Evidence from WASDE reports. *Journal of Agricultural and Applied Economics, 40*(1379-2016-112694), 89-103.

Isengildina-Massa, O., Karali, B., & Irwin, S. H. (2020). Can private forecasters beat the USDA? analysis of relative accuracy of crop acreage and production forecasts. *Journal of Agricultural and Applied Economics, 52*(4), 545-561.

Jaba, E., Robu, I., & Balan, C. B. (2017). Panel data analysis applied in financial performance assessment. *Revista Română De Statistică, 65*(2), 3-20.

Just, R. E. (1983). The impact of less data on the agricultural economy and society. *American Journal of Agricultural Economics, 65*(5), 872-881.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision making under risk. *Econometrica, 47*(2).

Levy, J. S. (2003). Applications of prospect theory to political science. *Synthese*, *135*(2), 215-241.

Lorton, S., & White, D. (2002). *Merchant's edge*. Champaign, IL: Stripes Publishing.

McDermott, R. (2004). Prospect theory in political science: Gains and losses from the first decade. *Political Psychology, 25*(2), 289-312.

Mercer, J. (2005). Prospect theory and political science. *Annu.Rev.Polit.Sci., 8,* 1-21.

Nofsinger, J. R. (2017). *The psychology of investing* (6th ed.). Milton: Routledge Ltd - M.U.A. doi:10.4324/9781315230856 Retrieved from https://www.taylorfrancis.com/books/9781351863315

Nofsinger, J. R. (2018). Behavioral aspects of commodity markets. In K. H. Baker, G. Filbeck & J. H. Harris (Eds.), () Oxford University Press.

Perloff, J. M., & Rausser, G. C. (1983). The effect of asymmetrically held information and market power in agricultural markets. *American Journal of Agricultural Economics, 65*(2), 366-372.

Shleifer, A. (2000). *Inefficient markets: An introduction to behavioural finance* OUP Oxford.

United states department of agriculture FY 2021 budget summary. .Government Printing Office. Retrieved from https://www.usda.gov/sites/default/files/documents/usda-fy2021-budget-summary.pdf

Von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, New Jersey: Princeton University Press.

Wang, Y., Zhu, Y., Zhang, S., & Wang, Y. (2018). What could promote farmers to replace chemical fertilizers with organic fertilizers? *Journal of Cleaner Production, 199,* 882-890.

Xie, R., Isengildina-Massa, O., Dwyer, G. P., & Sharp, J. L. (2016). The impact of public and semi-public information on cotton futures market. *Applied Economics, 48*(36), 3416-3431.

Can Any One Trader Beat the Market Consistently and How Is Performance Related to Gender and Experience

Abstract

Using private commercial grain trader data from a private company for the period from July 2011 to February 2018, this paper sheds light on the repeated performance of commercial grain traders purchasing corn and soybean meal for feed mills. Two time horizons were chosen to evaluate the commercial grain traders, monthly and yearly. The traders are first evaluated by their ability to repeat their performance over two consecutive periods with the use of a Fisher Exact ranking test. Further, the commercial grain traders were evaluated by the magnitude of profitability. To assess the difference in the profitability of traders a top and bottom performance test was used for each time horizon. We find that traders may be able to repeat their performance on the quarterly time horizon for the shorter time horizon we also found a difference in magnitude of profitability for the top and bottom-ranked trader. Additionally, we investigated the role of gender and experience on commercial grain trader performance.

Keywords: agricultural economics; futures markets; predictability; gender differences

Introduction

In order to guide policy decisions affecting agricultural markets, it is important to understand how those markets function. The behavior of market participants is one of the most important factors in the functioning of markets. Among others, previous studies on fair market participation for farmers (Hendrickson & Petty, 1994), or the role of speculators in agricultural commodity markets (Koziol & Treuter, 2019) have been paving the way to gain a deeper understanding of the role of agricultural market participants. This study uses a unique dataset that allows us to take a close look at the behavior and performance of commercial grain traders. We are particularly interested if certain traders are able to consistently perform better than others, and what role gender and experience play in the commercial trader performance.

Institutional Trading Background

This research focuses on an agribusiness firm that purchases large quantities of corn and soybean meal throughout the year. As such, the corn and soybean meal purchases are important input costs in this process, and the price at which they are purchased can have a material impact on the firm's profit margins. To effectively manage the price risk associated with these inputs, the firm employs a dedicated team of traders to hedge their costs and effectively lock in the purchase price of the grain prior to its physical delivery. This price risk management task is divided between two different groups of traders. One group is tasked with timing the purchase of futures contracts (long/bought futures positions), and the other group of traders – who are the focus of this study – are tasked with timing when to lock in the basis (long/bought basis positions) component of the hedge. Basis is formally defined as the difference between the cash price of a commodity and its respective futures price ($B_t = CP_t - FP$).

To illustrate, consider the general case of a firm that is hedging its future input purchases and is faced with upside price risk prior to establishing/contracting a price for the inputs. More formally, the cash price for a commodity observed in time *t+1* in market *j* is a stochastic random variable upon which grain traders for the hedging firm will form expectations about at time *t*, so that:

$$CP_{t+1,j} = E_t\left(\frac{CP_{t+1,j}}{\Omega_t}\right) + \left(U_{t+1,j}\right) \tag{1}$$

where $E_t(CP_{t+1,j}/\Omega_t)$ is the expectation of the future cash price formed in period t, and Ω_t denotes the information set available in period *t*, and $U_{t+1,j}$ is an expectations error. Our analysis assumes that the quantities of a commodity that will be purchased in period t*+1* are known at time *t*. In other words, any quantity risk that may exist as a component of overall input cost risk is ignored. This is not an unreasonable assumption in the case of our hedging firm, which has detailed knowledge of the amount of gain needed at its various feed mills (represented as market location *j*) at different times in the year.

In our case, to mitigate the price risk associated with $CP_{t+1,j}$ *o*ur hedging firm traders can use forward contracts, futures contracts, and basis contracts. Forward contracts, entered into in period *t*, establish a fixed cash price for period *t+1*. Traders can enter into these contracts with grain merchandising firms (elevators) and/or farmers. Hedged positions, formed at time *t* using either forward contracts or futures contracts tied to basis contracts, remove all of the conditional price risk. In other words, the conditional variance of $CP_{t+1,j}$ at time *t* is zero once these hedged positions are established. However, the decision of when to establish the hedged positions and lock in the price is complicated, hence the need for professional traders with experience and knowledge of market supply and demand conditions and what would be considered a good value to lock in effective purchase prices.

In the case of forwards contracts with resellers/merchandisers/elevators, the traders choose the timing of when to enter these contracts and lock in a price. Although there is some negotiation on the forward contract price, essentially, it is the elevator that offers the price, and this price is dictated or determined by supply and demand in the local cash market. With respect to the farmer-based forward contracts, our hedging firm offers the forward price – again subject to market conditions – and the farmer chooses the timing of when to contract for any time greater than six weeks prior to physical delivery. Besides forward contracts, farmers are also offered basis contracts where the basis is fixed and the futures price is determined later, or Hedge-to-arrive contracts where the futures price is set and the basis is determined later. Given that the marketing/risk management decision rests with farmers in the second case, all farmer contracting observations are excluded from the analysis in this first essay. However, farmer contract observations are included in the analysis presented in the third essay. In the context of equation (1), if traders' price expectations, $E_t(CP_{t+1,j}/\Omega_t)$, exceed the observed forward contract price, CFP_t^{t+1} for delivery in period *t+1*, and observed in period *t*, there will be an incentive to enter into a forward contract to lock in at least some amount of the anticipated grain that will be needed in period *t+1* for market location *j*.

Turning to the case when our hedging firm uses futures and basis contracts to manage price risk, the market timing and risk management decisions are separated in terms of type of contract and trader. The vast majority of trades in our sample data are basis contracts. Large long futures positions are accumulated over time to cover anticipated physical cash purchases, which have not already been forward contracted. Analogously to the forward contracting decision, if traders' price expectations, $E_t(F_{t+1}^{t+1}/\Omega_t)$, exceed the observed future contract price, F_t^{t+1}, in period *t*, there will be an incentive to buy futures contracts to cover at least some amount of anticipated future grain purchases in period *t+1* across market locations. The timing of these futures contract purchases in terms of quantity and pricing targets are made by a group of expienced traders. These trades are not part of our analysis. Instead, we focus attention on the risk management decisions of the group of traders who are responsible for buying basis contracts to help hedge the price risk associated with the future purchases of cash grain. When traders' basis expectations at time *t*, $E_t(B_{t+1}/\Omega_t)$, exceed the observed basis contract value B_t^{t+1}, which locks in the basis for delivery period *t+1*, there is an incentive to enter into the basis contract. This is because the lower the basis value that is locked in, the lower the effective hedged purchase price. This is illustrated by considering how the price risk is associated with the expected cash price from equation (1), $E_t(CP_{t+1,j}/\Omega_t)$, is hedged using futures and basis contracts. The effective cash purchase price $ECP_{t+1,j}$ in period *t+1* for market location *j*, can be locked in or established in period *t* by forming a hedge with long futures in period t:

$$\widehat{ECP}_{t+1,j} = FP_t^{t+1} + E_t(B_{t+1}) \tag{2}$$

Note that the expected basis, $E_t(B_{t+1})$, is a stochastic random variable, so hedging with long futures alone establishes the effective cash purchase price, subject to basis risk. In other words, because of basis risk the effective cash purchase price, $\widehat{ECP}_{t+1,j}$ is a stochastic random variable. However, if the basis is also locked in through the purchase of a basis contract in period *t*, all price risk is removed, and the effective cash purchase price is fixed:

$$ECP_{t+1,j} = FP_t^{t+1} + B_t^{t+1} \tag{3}$$

Note that if $B_t^{t+1} < E_t(B_{t+1})$, then $ECP_{t+1,j} < \widehat{ECP}_{t+1,j}$, so the lower is B_t^{t+1} in comparison to $E_t(B_{t+1})$, the lower is the effective cash purchase price. With this in mind, if traders can consistently buy basis contracts at levels below the resulting basis levels at the physical delivery time of commodity purchases, effective cash purchases will be cheaper, and price risk will be removed. In our sample data, the vast majority of transactions are basis contracts. The natural metric with which to evaluate traders' performance in terms of basis trading is the difference between the actual basis at physical delivery time *t+1*, denoted as B_{t+1}, and the basis established through a basis contract in period *t*, denoted as B_t^{t+1} We refer to this as the traders' margin:

$$TM = B_{t+1} - B_t^{t+1} \tag{4}$$

Note *TM* is positive when traders lock in basis levels – using basis contracts – at levels below basis at physical commodity delivery time. All of our subsequent empirical analysis is focused on *TM* for individual traders. The *TM* that we evaluate is generated as an internal metric by the firm and accounts for spatial differences in transportation costs across market locations. The firm's adjustment for transportation costs makes *TM* comparable across market locations and by the trader. It should also be noted that even if a trader locks in the cash purchase price of the commodity through a forward cash contract, the transaction is also evaluated in terms of *TM*. The forward contract price established at time *t* is converted into an equivalent basis value, B_t^{t+1}, using the difference between the cash forward contract price CFP_t^{t+1} and the current futures price F_t^{t+1}.

Using *TM* as a trader performance metric has certain drawbacks. For example, traders' decisions on when to lock in basis and prices may not be just based upon comparisons between their basis expectations and current basis contract values. They may also consider the conditional variance of basis and price – or in other words, the degree of price risk when making risk management decisions. If current market conditions are volatile – and price/basis risk high – locking in price and basis with hedges may be preferable, even if current basis contract levels are higher than the expected basis level at physical delivery time. In addition, although our firm tracks *TM* by trader, we do not know if it is used in any formal evaluation process, and individual traders may be unaware of their own performance in terms of *TM*. As such, their risk management decisions may be only indirectly influenced by their *TM* performance. However, *TM* is an interesting metric to evaluate trader performance and is perhaps the best available measure of whether traders can consistently lower the firm's effective purchase prices while also removing price risk. *TM* also has interesting implications with respect to market trading efficiency (EMH) in physical cash markets for corn and soybean meal. EMH implies that the future price at any point in time is equal to the markets' expected futures price at contract maturity, or more formally, $E_t F_{t+1}^{t+1} = F_t^{t+1}$. Therefore, on average, the margin ($F_{t+1}^{t+1} - F_t^{t+1}$) associated with any trading position taken at time *t* by traders/speculators would be zero. Unbiased and efficient future prices, which adjust quickly to current market information, do not provide systematic positive trading margins. Analogously, if physical cash basis markets are also EMH unbiased ($E_t B_{t+1}^{t+1} = B_t^{t+1}$), then taking basis contract positions at time *t* should not result in systematic positive trading margins. However, little academic research has investigated if this is true, and it is ultimately an empirical question. McKenzie, Isbell, and Brorsen (2019), is one of few studies that have addressed this issue, found that forward basis contracts for the delivery of grain at the Gulf contained systematic pricing biases or risk premiums.

Literature

Aulerich, Irwin, and Garcia (2013) compared individual non-commercial traders' performance in agricultural futures markets, specifically corn, live cattle, and coffee futures markets. They investigated whether non-commercial traders can be successful repeatedly. Non-commercial traders are also referred to as speculators. Their study focuses on a rather large set of traders from the Commodity Futures Trading Commissions Large Trader Reporting System and focuses on corn, cattle, and coffee traders. Besides different commodities, they chose different time horizons, monthly, quarterly, and yearly. Aulerich et al. (2013) used two main tests for their analysis: a winner and loser rank test and a top to bottom decile test. The profitability measure in this study is based on daily aggregate profits. Their study revealed persistence in the performance of traders in terms of making profits.

Luck is the main factor determining trader performance, according to Hartzmark (1991), who tested if futures traders in agricultural commodities can consistently earn positive profits. The data used consisted of commercial and non-commercial traders in nine different agricultural futures markets. When tested for forecast ability, only commercial traders in the pork belly markets were on average able to show significant forecasting abilities. Commercial traders were also credited with a slightly superior forecast ability than the speculators, which may be due to better and timelier information access. Hartzmark also tested the skill hypothesis where traders' performances are compared in two time periods. In general, he found no support for non-commercial traders' performance to be persistent across time.

Leuthold, Garcia, & Lu (1994a) extended Hartzmark's results with respect to traders' outstanding forecast performance in the pork bellies market. They also found that a subset of elite traders, as they call them, can forecast price movements in the market and profit from their forecasts. Leuthold et al. suggest that specific traders can increase their skill and ability to forecast market movement levels with accumulated experience and market knowledge. Other traders, especially short-hedgers and spreaders, did not show a superior forecasting ability when trading, which could be due to their specific risk management strategy use of the futures market. For example, those traders may simply be short hedging their cash market positions, with no motivation to profit from their futures trades. In our case, futures market risk management hedges are conducted by a separate set of traders, whose main focus are futures market transactions. The traders our research focuses on are covering the demand of the feed mills while using forward contracts and basis instead of the commodity price. While our traders are not short hedging themselves, they may find themselves in situations at times where they need to buy regardless of the basis or price to keep their feed mills from loss of production. For this reason, to focus on the forecasting ability of the traders, short notice spot market transactions are not included in this research.

Numerous studies have compared the risk-taking behavior of males and females in diverse contexts. A meta-analysis from 1999 about gender differences in risk-taking (Byrnes, Miller, & Schafer, 1999) concluded that generally, men are more likely to take risks than females but to a variable extent across ages and contexts. Croson & Gneezy (2009) confirm the finding of women being more risk-averse than men in a majority of tests. They also provide speculative reasons as to why women might make more conservative choices when faced with risk. Firstly, adverse outcomes are feared more strongly by women who are also more prone to feeling fear or nervousness. Women experience emotions more intensely than men, which might affect the utility of a risky choice. Secondly, women show less overconfidence in their investment decisions. Women evaluate their success less favorably than men would do in the same situation, leading to men more readily accepting risky gambles. Lastly, men are more drawn to ego-involving situations, which women in return might perceive as a threat and shy away from, given their different motivations. The overconfidence problem mentioned by Croson and Gneezy is also named as a cause of high levels of financial trading by Barber & Odean (2001). They state that humans who are overconfident about their abilities, knowledge, and prospects trade more frequently. Barber and Odean (2001) explain that rational investors only invest time and information in trading when they can increase their expected utility. Simultaneously, overconfident investors invest too much in gaining information and lower their expected utility by executing too many trades. The greater overconfidence that psychologists have found in men leads to men trading more than women, which leads to men suffering lower performance due to the increased trading rate. Barber and Odean (2001) found that single men reduced their returns by 1.44 percent compared to single women by overly frequent trading. A lower willingness to take risks in personal financial contexts could result in women opting for low-risk portfolios. And selecting low-risk portfolios will lead to lower expected returns over long periods compared to higher-risk portfolios often chosen by men leading to a gap in wealth (Nofsinger, 2017). Across the professionals employed in finance, gender differences are often nonexistent. Atkinson, Baird, & Frye (2003) compared the performance of fixed income funds managed by men and women with similar education and found similar investment behavior. Dwyer, Gilkeson, & List (2002) Investigated mutual funds investors' behavior and found no significant difference for female investors with financial and investment knowledge when taking investment risks. They further conclude that knowledge disparities can partly explain the differences in women and men making financial decisions.

Experience is another factor affecting trading behavior according to the behavioral finance literature. Previous studies have mainly focused on stock and options markets (Barber & Odean, 2001; Hoffmann, Post, & Pennings, 2015a) rather than commodity markets. However, all share similar findings, with less experienced traders exhibiting a more unsatisfactory performance in terms of returns than more experienced traders do. That traders improve in their investments over time was also found by (Nicolosi, Peng, & Zhu, 2009), who found that retail investors in the stock market can improve their trading with increased experience. Stock investors not only increased portfolio returns with increased experience but also improved the quality of their trades.

Data

The dataset for this research was provided by a company purchasing corn and soybean meal for further processing for livestock feed. This dataset comprises soybean meal and corn trades executed between July 2011 and February 2018. In total, there are more than 30,000 individual trades or observations executed by 13 traders. Out of the 13 traders, four are females and nine males, with their grain trading experience ranging from beginner level to over 30 years in 2018. Besides information about the trader, the transactions contain information about the transaction dates, locations, transportation mode, prices, etc. In the dataset are 70 different locations that are destinations for commodity deliveries. For a trade to be labeled as a forward contract and to be included in this data analysis, 12 days or more have to lie between the contract agreement and the delivery date. All trades in the dataset based on the farmers' decision to sell their products are excluded. Only transactions initiated by the company traders are used for the data analysis. Most of these transactions are basis contracts between the purchasing company and large resellers or merchandising firms. Several conversions of volumes and prices had to be conducted to unify the units to U.S. dollars per bushel. These conversions were necessary to achieve comparability of the transactions of the two commodities. All amounts that were provided initially in lbs and were converted to bushels. Soybean meal prices, and basis were in U.S. dollars per short ton, and for the corn trades, numbers were already in U.S. dollars per bushel. Soybean meal was converted with 2000 lbs in one short ton and 54 lbs soybean meal per bushel, while 56 lbs of corn are equal to one bushel (Common weights.). Ultimately, all *TM*s for corn and soybean meal were converted into dollars per bushel and the two types of commodity trades were combined in the analysis.

All following tests are conducted with a merged dataset consisting of both corn and soybean meal trades. The two commodities' *TM*s were compared in a paired t-test that did not significantly differentiate the two commodities. Unfortunately, corn and the soybean meal monthly data set by themselves did not contain a sufficient amount of traders or observations to analyze by a single commodity. Not all traders are active simultaneously, meaning, in this case, they don't have deliveries arriving in a respective time horizon. Most traders specialize in one commodity and trade mostly just either corn or soybean meal, which can lead to periods, especially short term, where only three traders receive orders of either commodity. Additionally, some of the traders were hired after the first day of the dataset, which led to increasing trader numbers in the later periods.

The most important variable for this analysis is the *TM* per bushel in U.S. dollars, which is calculated for every trade. It is the difference between the previously agreed-upon delivery price and the spot market price at the delivery time. The *TM* quantifies in retrospect how good or bad the forward agreement was. It expresses how much was over or underpaid compared to a spot market purchase at delivery. If the *TM* is 0, the prearranged price and the actual spot market price end up equal. A negative *TM* would indicate that the prearranged forward contract price was higher than the final spot price observed at the forward contract's maturity. A negative value for the *TM* represents an over-payment. A *TM* greater than 0 would indicate that the forward contract price was less than the final spot market price observed. A positive *TM* means that the trader has achieved a lower price than if it was purchased in the spot market on the day of delivery. Table 1 shows the mean *TM* for all trades per trader over the whole dataset.

All transactions are evaluated based on the delivery date for this specific analysis. Which period a trade falls into is determined by the delivery date. The grouping of trades by delivery date is essential as all trades have a different contract length, and it can be assumed that the demand for the commodities is steady throughout the year. Traders might set up more transactions at a particular time, but the different expected delivery periods make it difficult to compare them. The data is most comparable by grouping the trades by the delivery date. *TM*s are calculated for every transaction per trader per quarter or year based on the month in which the delivery took place. A descriptive table of the yearly mean *TM*s per trader is displayed in table 2. In table 3, the quarterly *TM*s per trader are displayed. The year's quarterly periods are separated by the crop production cycle, starting with an assumed harvest on September 1st. Accordingly, the first quarter comprises the months of September, October, and November. The second quarter includes December, January, and February. Next, the third quarter comprises March, April, and May, while the last quarter covers June, July, and August.

Methods

Four main tests were used for the analysis of this study. First, a pooled winner and loser rank test was applied, second, a top and bottom performance test, third, a gender performance test, and lastly, an experience test. All data analysis was conducted with SAS (Insititue Inc., 2019)[1] SAS and (Microsoft) Microsoft Excel software.

Pooled Winner and Loser Rank Test

The pooled winner and loser rank test is conducted for yearly and quarterly time horizons. For this nonparametric two-way winner and loser contingency table test, traders are categorized into winner and loser categories. All traders are ranked from best to worst performance for each period by their average *TM* of forward contracts scheduled for delivery in the respective time frame. The ranking ranks the trader with the highest mean *TM* in the first and lowest on the last rank. The test uncovers consistent or persistent superior trading performance over time. In other words, we are able to determine if some traders exhibit superior trading skills over time.

The analysis starts with creating sets of adjacent periods, t and t+1. In the case of a yearly time horizon, 2012 and 2013 represent the first set of adjoining periods, and 2016 and 2017 are the last. In total, there are five adjacent pairs of periods. For the quarterly data, the first period starts on December 1st of 2011, and the last period ends on February 28th, 2018, yielding 20 usable period pairs. To be included in a bracket of adjoining periods, traders have to be active in

[1] The data analysis for this paper was generated using SAS software. Copyright © [2002-2012] SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA.

both. If a trader was active in t but not in t+1, they are not included in the pair of periods and disregarded. Periods in which less than four traders have been active are not considered and excluded from the analysis. Given the limited number of traders in this data set, the top two performing traders are categorized as winners and the bottom two as losers. Most adjacent periods have at least four active traders, which would provide two winners and two losers. If more than four traders are active, only the top two and bottom two are winners or losers, and the other performances of traders are seen as mediocre, neither a winner nor a loser.

For the winner and loser rank test, we count how many traders fall into one of the following categories per period pair. A count is recorded in category A of the number of traders who have been a winner in period t and are winners again in period t+1. The second count, category B, is the number of traders who have been a winner in period t but fell into the loser category in t+1. The third category, C, includes the counts of traders who have been a loser in t but switched to the winner category in the following period t+1. Lastly, category D, the fourth group of counts, are all traders who were losers in period t and again in the subsequent period t+1. In every period, t are two winners and two losers, but it is possible that the total count of traders in t+1 is less than four. When the total number of placings in t+1 is less than four, the period pair needs to have more than four traders, and one trader, who was either a winner or loser in t, falls into the middle range (neither winner nor loser) in t+1.

After determining how many of the four traders of the first period (t) in each bracket are either winners or losers in the second period (t+1), all counts for categories A, B, C, and D are summed up to be pooled.

A Fisher's Exact Test is applied to the pooled data to test for the traders' persistent placement. The Fisher's exact test is used in cases where the expected number in one or more cells is too small for a chi-square test. Like the chi-square test, the test evaluates whether distributions of the average *TM*s of traders categorized in A, B, C, or D are the same. In our case, the four categories are traders in who are winners in t and t+1, winners in t and losers in t+1, losers in t and winners in t+1, and losers in t and t+1. The 2 by 2 table has two dimensions: the rows, the columns, and in a random distribution, a conditional probability of 50% would be expected between the two classifications. In our case, the null hypothesis is that the trader ranking in t+1 is independent of the ranking in t. A rejection of the null hypothesis would indicate that the ranking of a trader in t+1 is dependent on the trader's ranking in t, and this would be evidence that trader performance was persistent across time. In other words, traders who exhibited superior performance in t were more likely to also have superior trading performance in period t+1.On the other hand, traders who exhibited inferior trading in t were also more likely have inferior trading performance in t+1. These findings would be consistent with the notion that some traders have consistently superior or inferior trading skills.

Top and Bottom Performance Test

Given that superior trading skills may only occur at the extremes of trading performance, we also use a Top and Bottom performance test to measure the magnitude of the differences in the top and bottom traders' *TM*s. As in the previous test, time brackets of adjacent time periods are created of traders active in both periods. Again, the test is designed to uncover superior trading performance by determining if trading margins are significantly different over time between traders who initially are ranked as best and worst performers.

First, traders are ranked by their performance in t—the highest average *TM* ranks in the first spot and lowest average *TM* in the bottom spot. Next, the trader's average *TM* is calculated in t+1, and the average *TM* from t+1 is attached to their ranking in t. Their performance in t still ranks the traders, but the values of the average *TM*s attached are from t+1.

Two comparisons are conducted to assess the magnitude of difference between the top and bottom rankings. In the first comparison, the top-performing spot is compared to the bottom performance. The second comparison compares the average of the two top performances to the two bottom performances' average. All average *TM*s are collected for the top performing traders and the respective bottom performing traders. This step is necessary to use paired t-tests to assess if the difference between the winner and loser group is different from zero.

The t-test is recommended for top and bottom rank tests by (Carpenter & Lynch, 1999), who compared mutual funds' performance persistence test methodologies. If a normal distribution or independent observations assumption fails to hold for the t-test, the Wilcoxson Signed Rank test provides a viable alternative. The Wilcoxon Signed Rank Test is conducted by Aulerich et al. (2013) in addition to the paired t-test.

Calculations for the Wilcoxon Signed Rank test are carried out by first defining the null hypothesis, in this case, that the median difference (M) between the top (a) and bottom (b) group is equal to zero. Secondly, each paired difference (d) is calculated for the observation pairs of a_{ti} and b_{ti}. $d_{ti}= a_{ti} - b_{ti}$. Next, the d_{ti}s are sorted by rank while ignoring the sign. Then each rank is labeled with its sign, according to the sign of d_{ti}. In the next step, all positive d_{ti}s are summed up to $W^+$and all negative d_{ti}s are summed up to calculate W^-. Choose $W = \min(W^-, W^+)$, and with the help of a table, the p-values can be obtained.

If the *TM* difference in t+1 between top and bottom performing traders is significantly different from zero, persistence in the trader performance can be assumed. In other words, strong trading performance in t should be carried over to t+1 if certain traders truly have consistent and superior trading skill. If the average *TM* difference between top and bottom performing traders declines between t and t+1, and is not significantly different from zero, this would indicate trading performance is random and cannot be attributed to skill. Therefore, the H_0 of the t-tests and the Wilcoxon rank test states that the top and bottom trader's performance is not different than 0.

The Top and Bottom Performance test was applied to the same corn and soybean meal transactions dataset as the previous rank rest. In addition, the analysis was applied for the yearly and quarterly periods.

Gender Performance Test

The third part of the data analysis focuses on the performance of the different genders. Here we compare the average yearly *TM* and the achieved by male and female traders in 2017 and over the whole dataset with an independent two-sample t-test. The year 20107 was chosen as this year had the most active traders. Also, focusing on one year has the advantage that all mean *TM*s were achieved under similar market conditions. As for the previous tests, all transactions are based on their delivery date that falls into 2017.

The SAS command for the two-sample t-test also provides the folded F test results to test for the assumption of homogeneity of variance. SAS uses F for the two-sample F-test, which is a ratio of sample variances; $F = s_1^2/s_2^2$. For the two-sample t-test, it is arbitrary which sample is labeled 1. For the folded F-test SAS chooses the larger sample variance to be placed in the numerator. The folded F-Test is calculated as follows (UCLA: Statistical Consulting Group):

$$F' = \frac{\max\left(s_m^2, s_f^2\right)}{\min\left(s_m^2, s_f^2\right)} \tag{5}$$

Where s_m^2 is the male sample variance and s_f^2 is the female sample variance. The folded F-test tests the H0: The variances of females are equal to the variances of males. The SAS t-test provides a pooled t-test appropriate for equal variances if we accept the null hypothesis. Suppose the assumption of homogeneity of variances is violated, and we have to accept the alternative hypothesis for the folded F-test. In that case, the SAS t-test provides the Satterthwaite alternative to the pooled method. The pooled t-test statistic is calculated as (Kent State University Libraries, 2021):

$$t = \frac{\bar{x}_f - \bar{x}_m}{Sp\sqrt{\frac{1}{n_f} + \frac{1}{n_m}}} \tag{6}$$

with:

$$Sp = \sqrt{\frac{\left(n_f - 1\right)s_f^2 + (n_m - 1)s_m^2}{n_f + n_m - 2}} \tag{7}$$

Where $\bar{x}$ is the mean of the respective gender, m indicating male and f female. The sample size is n, s is the standard deviation, and Sp is the pooled standard deviation. The calculated t value can then be compared to a t table critical t value and with degrees of freedom $df = n_f + n_m - 2$.

When the folded F test does not indicate homogeneity of variance between males and females, the Satterthwaite t value is calculated as:

$$t = \frac{\bar{x}_f - \bar{x}_m}{\sqrt{\frac{s_f^2}{n_f} + \frac{s_m^2}{n_m}}} \tag{8}$$

Here $\bar{x}$ is the mean of the gender, with f standing for female and m for male. The sample size n and standard deviation s complete the equation. The t value calculated with this equation is then compared to a critical t value from a table. The degrees of freedom are calculated as follows:

$$df = \frac{\frac{s_f^2}{n_f} + \frac{s_m^2}{n_m}}{\frac{1}{n_f - 1}\left(\frac{s_f^2}{n_f}\right)^2 + \frac{1}{n_m - 1}\left(\frac{s_m^2}{n_m}\right)^2} \tag{9}$$

Experience Test

The fourth and final test focuses on the experience of the traders, with the question being does level of trading experience influence trading performance. Analogous to the gender performance test, we focus this test on one single year, 2017. Limiting the data to one year provides the same market conditions for all traders and is most suitable to compare experience levels. Due to the limited number of active traders each year, 2017 was chosen as the year with the most active traders. To analyze if the trader's experience level influences the average *TM*, we split the traders into two different experience levels. The ten traders of 2017 are split into experience levels of three years or less and four years and more, based on their work experience in 2017.

Results

The results section will first look at the winner and loser rank test results and at the top and bottom performance tests. Next, the Results are split up into quarterly time periods and yearly time periods.

Quarterly Results

The descriptive statistics for the quarterly average mean *TM*s per trader are shown in table 3, and table 4 shows the more detailed mean *TM*s per quarter. The ranking in table 3 shows that trader 2 was most successful over the whole time span of the dataset by quarters. Trader 2 achieved an average mean *TM* of 0.36 U.S. dollars per bushel. On the bottom end, trader 6 averaged a mean *TM* of 0, None of the traders had a negative average mean *TM* over all the quarterly periods.

Quarterly Winner and Loser Rank Test Results

The winner and loser distribution in t+1 based on their ranking in t for the quarterly data is displayed in table 5. The pooled count of the 20 quarterly brackets is recorded in table 6. Out of 40 winners in t based on their average *TM* performance, 23 traders achieved a winning ranking again in t+1. Therefore, the conditional probability of being a winner in t and again in t+1 is 58%. Only 12 out of the 40 winners in t slipped into the loser category in the next quarter, equaling a 30% conditional probability of being a winner in t and a loser in t+1. Those 40 traders in the loser category in t had a conditional probability of being a loser again in t+1 of 60%. Thirteen out of the 40 traders in the loser category in t could flip their ranking to a winning placing in t+1. The conditional probability of being in the losing category in t and the winning category in t+1 is 33%

Fisher's exact test results for the quarterly pooled data in Table 7 show that the two-sided p-value is 0.0178. Therefore, we reject the null hypothesis assumption of independence of trader performance across periods at the 5% significance level. This result indicates that based on our quarterly results, there is a group of superior traders that consistently outperform other traders. It would appear that some traders have superior or inferior trading skills.

Quarterly Top and Bottom Performance Test Results

Table 8 shows the *TM*s of the top and bottom-ranked traders in the respective quarters. The top and bottom traders in t achieve an average mean *TM* of 0.37 U.S. dollars per bushel and 0.16 U.S. dollars per bushel, respectively in t+1. The top trader outperforms the bottom trader in t+1 by 0.21 U.S. dollars on average. When comparing the top and bottom two traders observed in t by their combined average mean *TM* per bushel in U.S. dollars, the results show that the difference between those two groups is 0.04 U.S. dollars per bushel. The two top traders observed in t achieve an average mean *TM* of 0.29 U.S. dollars per bushel in t+1, while the two bottom traders observed in t achieve a mean *TM* of 0.25 U.S. dollars per bushel in t+1. Turning to the statistical comparison in table 9, the top and bottom two traders' performances in t+1 reveal that the quarterly time horizon performance does differ statistically from zero. The top and bottom trader's 0.21 U.S. dollar *TM* difference in t+1 is significant at the 0.05% level for the quarterly data. The t-test has a p-value of 0.0146 and the Wilcoxon test has a p-value of 0.0182. The quarterly comparison of the top and bottom two combined average mean *TM*s per U.S.

dollars per bushel in t+1 shows mixed evidence of a statistical difference. The t-test indicates the *TM* difference is significant at the 10% but not at the 5% level (p-value 0.0529), while the Wilcoxon p-value of 0.125 indicates no statistical difference at the 10% significance level.

Summary Quarterly Results

The results of the quarterly data show that the ranking of the traders by the mean *TM* is not random. But the magnitude of the differences is only significant when comparing the top and bottom one trader when comparing the performance, measured in the mean *TM*, in two consecutive quarters. The difference is insignificant when comparing the top two trader's combined average mean *TM* to the bottom two trader's combined average mean *TM* in t+1. This would suggest, superior trading skill is only evident at the extreme tail of the trader performance distribution. The descriptive statistics show that the average mean *TM* was positive for all traders besides trader 6, who averaged a 0 average mean *TM*. Five out of the 14 traders in the dataset achieved an average mean *TM* of over 0.2 U.S. dollars for the quarters in which they were receiving purchases.

Yearly Results

Table 2 shows the average mean results for the descriptive statistics of the yearly data, and table 10 shows more detailed results of the yearly mean *TM*s. When looking at the average mean *TM*s of the years when the traders received orders, the table reveals that trader 12 is the most profitable trader with an average mean *TM* of 0.36 U.S. dollars per bushel per year. On the other hand, trader 6 ranked in the bottom position with a slightly negative average mean *TM* of -0.01 U.S. dollars per bushel.

Yearly Winner and Loser Rank Test Results

The pooled yearly winner and loser rank test for the ranking of traders by mean *TM*s comprises five-period brackets from 2012 to 2017. Table 11 shows the distribution of winners and losers in t and t+1 per two-year brackets with the total number of active traders in the two-year bracket. The pooled counts of the placements from Table 11 can be found in Table 12. The pooled table reveals that three out of the ten traders who were winners in t were again winners in t+1, revealing a conditional probability of a winner being a repeat winner of 30%. More winners in t were turned into losers in t+1 than winners being able to repeat their success. Out of the ten winners in t, seven turned into losers in t+1, representing a 70% conditional probability of being a winner in t and a loser in t+1. Out of the total ten losers in t, six turned into winners in t+1. The conditional probability of losers in t becoming winners in t+1 is 60%. Out of ten initial losers in t, three could not improve their ranking. They stayed in the losing category, which means the conditional probability of being a loser in t and repeating the bottom ranking is 30%.

The pooled counts of the rankings for the yearly data do not look to be evenly distributed at first. The Fisher's exact test revealed a two-sided p-value of 0.1789, as Table 13 shows. At the 5% significance level, the Fisher's exact test H_0 of a random ranking of the traders' performance cannot be rejected. The yearly results, unlike the quarterly results, indicate no consistency in superior trader performance.

Yearly Top and Bottom Performance Test Results

First, the top and bottom performance test calculates the mean *TM* per trader in t+1 while the traders are ranked based on their performance in t. Table 14 shows the mean *TM*s in t+1 for each top and bottom-ranked trader of t and the combined mean *TM*s of the top two and bottom two traders on a yearly basis. The comparison results of the top and bottom traders can be found in Table 15. The results table shows that the yearly average mean *TM* in t+1 for the top-ranked trader of t was 0.09 U.S. dollars per bushel. On the other side, the average mean *TM* in t+1 for the bottom-ranked trader in t was 0.33 U.S. dollars per bushel. This larger average mean *TM* means that a trader who ranked bottom in the first year was on average 0.24 U.S. dollars per bushel more profitable in t+1 than the top-ranked trader in t+1. The same can be observed for the yearly analysis when looking at the combined average mean *TM*s for the top two traders of t in t+1, and the combined average mean *TM* of the bottom two traders of t in t+1. This is evidence of performance reversal and indicates that superior performance is not consistent for traders.

While not as big of a difference as the top and bottom one trader, the bottom two ranked traders in t, achieve an average of 0.18 U.S. dollars per bushel more *TM* than the two traders ranked on top in t. The combined average mean *TM* in t+1 for the top two traders of t is 0.12 U.S. dollars per bushel and therefore higher than the average mean *TM* of 0.09 U.S. dollars per bushel in t+1 for the top one trader in t. This difference and the larger combined average mean *TM* of the top two traders compared to the top one trader in t indicate that the second-ranked trader in t outperforms the top trader of t in t+1. Comparing the top and bottom traders with the paired t-test and Wilcoxon signed-rank test reveals no significant difference in the performance of the top and bottom trader in t+1 at the 5% significance level. But both tests, the t-test p-value is 0.065, and the Wilcoxon p-value is 0.063, are significant at the 10% level. When comparing the performance in t+1 of the107 two top and bottom traders on a yearly level, the t-statistic value of 0.0529 is barely insignificant at the 5% significance level. Still, the Wilcoxon p-value also misses the significance level of 10%, as the p-value of 0.125 shows. The results make it difficult to decline the H_0 of no difference in trader performance in t+1 based on their ranking in t. While the yearly results do not indicate a difference in the achieved *TM* per bushel at the 5% level, it can't be denied that they do differ at slightly higher significance levels. However, importantly, given that we find evidence of performance reversal, our results indicate that traders are unable to display superior trading skills over longer periods.

Summary Yearly Results

The pooled winner and loser rank test results show that the traders' ranking is random for the yearly data. Also, the top and bottom performance test did not show a strong significance for the magnitude of the difference between the top and bottom traders. Similar to the quarterly results, the average mean *TM* per year ranks from -0.01 U.S. dollars per bushel at the bottom to 0.36 U.S. dollars per bushel per year at the top.

Comparison of Results

Comparing the results from the quarterly rankings to the yearly rankings of traders, we only find significance of trader ranking for the quarterly periods. These results show that the ranking of traders is random for the yearly data. While traders may repeat their ranking in two consecutive quarters, they may not be able to repeatedly rank in the same spot for more than two consecutive quarters repeatedly. When comparing the average mean *TM*s per bushel, the results show that for the quarterly data, the top trader in t has a significant advantage in t+1 over the bottom trader in t. But the results lose significance when even comparing the performance of the top two traders to the bottom two traders of t in t+1. For the more extended periods of the yearly data, the magnitude in the difference between the top and bottom traders in t is not significant in t+1. But even while not significant, the bottom traders from t now outperform the top trader in t+1. These results indicate that traders can not constantly outperform each other in the longer run. When comparing the descriptive statistics of mean *TM*s over the whole dataset (table 1), the numbers look different from the descriptive statistics for the yearly or quarterly average data. When looking at the single trades regardless of the time period, the top trader, trader 11, achieves a 0.27 U.S. dollars average *TM* per bushel per trade. This number is lower than the top average mean *TM* of either the yearly or quarterly data. Additionally, from all three descriptive statistics tables, the top-ranked traders have higher standard deviations than the lower-ranked traders, which may indicate higher risk-taking.

Gender Performance Results

First, it has to be noted that the small number of female traders to male traders in the dataset makes it hard to draw definitive conclusions about gender differences. Still, we found it interesting to compare traders' performance split by gender to the literature. For a fair comparison of the two types of traders, 2017 trades alone were compared. All trades that occurred in 2017 are accounted for, and they are not split into quarters. The number of trades done by females and males in 2017 was nearly equal, even though only three out of the ten traders were female. Females' total number of trades performed in 2017 was 2,980, and male traders made 2,819 trades. A more detailed division of the traders and their trades in 2017 can be found in table 16. The average *TM* of all trades set up by females is 0.17 U.S. dollars per bushel, while the average male *TM* was 0.09 U.S. dollars per bushel. On average, females achieved a 0.08 U.S. dollars better *TM* per bushel than their male colleagues. The folded F-test did not support the null hypothesis of homogeneity of variances. The rejection of the assumption of homogeneity of variances points to the Satterthwaite t statistic being the practical test to test if male and female performance difference is statistically significant. The p-value 0<.0001 of the Satterthwaite t-test indicates that the difference between the female and male transactions in 2017 is strongly significant, even at the 1% level. The results of the statistical analysis are presented in table 17.

However, when comparing gender differences in trading performance for the whole dataset, male traders are more successful in generating a higher average mean *TM* per bushel. Descriptive statistics over of the traders split by gender are in table 18. Over the whole period, there are 3,633 transactions initiated by female traders and 25,556 transactions initiated by male traders. Overall, the males achieved an average *TM* of 0.1879 U.S. dollars per bushel, and the female traders 0.1574 U.S. dollars per bushel. On average, the male traders are 0.0305 U.S. dollars more profitable than the female traders. The difference is significant at the 5% level, as shown in table 19.

Overall, our results provide mixed evidence as to differences in trading performance across gender, and the small number of female traders in our sample means it is not possible to generalize our results and make any definitive conclusions.

Influence of Experience on Trader Performance Results

To compare trader experience, the year 2017 was chosen to focus on all trades that occurred during a single year to account for best comparability. Traders were split into two groups based on their work experience in 2017. The first group comprises all six traders with three years or less of work experience in 2017. The second group had four years or more of work experience and consisted of four individual traders. A table of the traders with their experience can be found in table 20. The trader group with three years or less of work experience had 3,357 trades in 2017, and the more experienced trader group accumulated 2,442 trades in the same year. The less-experienced traders outperformed the more experienced traders by on average 0.101 U.S. dollars per bushel in 2017. While the more experienced group averaged 0.075 U.S. dollars per bushel in *TM*, the less experienced group achieved an average *TM* of 0.176 U.S. dollars per bushel. The folded F-test result led to the rejectiona of the assumption of homogeneity of variances for the two groups. With a p-value of <.0001, the p-value of the folded F-test pointed towards the use of the Satterthwaite estimation for the t-statistic. The Satterthwaite p-value of <.0001 shows that the difference between the less experienced and more experienced traders is strongly significant at the 1% significance level. Results of the statistical analysis for the experience level are in table 21.

As for the gender comparisons, it has to be highlighted that the small number of traders does not allow a generalization of our results. At the same time, it is possible that the results are random and would have turned out differently if a different year had been chosen. In our case, it may be possible that the less experienced traders took on more risky trades and were rewarded for their risk-taking. From table 17 it can be seen that the standard deviation of *TM* for the less experienced traders was higher (0.3175) than for the experienced traders (0.2868).

Discussion

The results of the winner and loser rank test show mixed signals. For the shorter term, the quarterly data shows persistent trader ranking. Whereas those results cannot be observed for the longer yearly time horizon.

Looking at the difference in magnitude of the top and bottom traders, the top and bottom performance test only show significance at the 5% level for the top and bottom trader comparison for the quarterly period. Here the top trader was on average 0.21 dollars per bushel more profitable than the bottom trader. The difference in magnitude for the yearly *TM*s is significantly different at the 10% level for the top and bottom traders, but in this case, the bottom trader outperforms the top trader by, on average, 0.24 dollars. This result does not support the idea of consistently superior trader performance, but instead shows evidence of performance reversal. The gender tests are strongly significant and show that the female traders outperform their male coworkers by, on average, 0.08 dollars per bushel. Unfortunately, the result is not generalizable due to the small number of traders. The generally accepted theory of overconfident male traders in behavioral finance (Barber & Odean, 2001) can not be tested here, as commercial purchases differ from stock market trading. At the same time, several studies find no significant difference in trading behavior for males and females who are professionals in the financial world (Atkinson et al., 2003; Dwyer et al., 2002). It may be possible that the female traders engaged in less risky foreward trades compared to the male trader, who may have suffered losses from engaging in riskier trades which may be in line with findings of women being more risk-averse (Byrnes et al., 1999; Croson & Gneezy, 2009; Nofsinger, 2017).

Our experience comparison revealed surprisingly that the less experienced group of traders outperformed the experienced traders on average by 0.10 dollars per bushel. While these results are not conclusive for the whole population of commercial grain traders, our results to be at odds with theory from the behavioral finance literature, which suggests trader experience is positively correlated with performance (Hoffmann, Post, & Pennings, 2015b; Nicolosi et al., 2009). However, it may be possible that the experienced traders exhibited less risky behavior when engaging in forward contracts and hence achieved smaller average *TM*s – a classic risk-return tradeoff.

Limitations and Recommendations

The study was performed with a dataset consisting of a small number of traders. It would be interesting to compare the performance of more grain merchants, possibly working for different companies. Further, it would be interesting to test the performance of grain traders for different commodities, as this dataset only included corn and soybean transactions. It may also be of interest to understand the performance of other market participants, such as the producers. Understanding the producer performance and the behavior leading to the performance would allow the creation of targeted programs to educate producers and create information tools to further the education of all market participants. Finally, the gender and experience comparisons were performed for only one year of trading activity due to the limited data. A longer time frame would be ideal.

Conclusion

While past performance can indicate future performance in the short term, these findings could not be observed for our study for the yearly time horizon. Both results from the gender and experience comparisons were unexpected, which shows that commercial traders in a risk-management setting cannot be directly compared to stock traders or non-commercial speculators. The results prove that further research is needed to better understand the behavior of all market participants in the agricultural field. With a better understanding of market participants behavior, government spending could be used more effectively to educate all market participants and ultimately make markets fairer and more efficient.

References

Atkinson, S. M., Baird, S. B., & Frye, M. B. (2003). Do female mutual fund managers manage differently? *Journal of Financial Research*, *26*(1), 1-18.

Aulerich, N. M., Irwin, S. H., & Garcia, P. (2013). Returns to individual traders in agricultural futures markets: Skill or luck? *Applied Economics, 45*(25), 3650-3666. doi:10.1080/00036846.2012.727979

Baer, J. B., 1881, & Saxon, O. G., 1892. (1949). *Commodity exchanges and futures trading: Principles and operating methods*. New York: Harper

Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics, 116*(1), 261-292. doi:10.1162/003355301556400

Black, F. (1976). The pricing of commodity contracts. *Journal of Financial Economics, 3*(1-2), 167-179.

Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin, 125*(3), 367-383. doi:10.1037/0033-2909.125.3.367

Carpenter, J. N., & Lynch, A. W. (1999). Survivorship bias and attrition effects in measures of performance persistence. *Journal of Financial Economics, 54*(3), 337-374. doi:10.1016/S0304-405X(99)00040-9

Chiappini, R., & Jégourel, Y. (2019). Explaining the role of commodity traders: A theoretical approach. *Economics Bulletin, 39*(3), 2002-2013.

Common weights. Retrieved from https://equi-analytical.com/common-weights/

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature, 47*(2), 448-474. doi:10.1257/jel.47.2.448

Dwyer, P. D., Gilkeson, J. H., & List, J. A. (2002). Gender differences in revealed risk taking: Evidence from mutual fund investors. *Economics Letters, 76*(2), 151-158. doi:10.1016/S0165-1765(02)00045-9

Fishe, R. P. H., & Smith, A. D. (2012). Identifying informed traders in futures markets. *Journal of Financial Markets,* (15), 329-359.

Geman, H. (2005). *Commodities and commodity derivates* John Whiley & Sons Ltd.

Grossman, S. J. (1977). The existence of futures markets, noisy rational expectations and informational externalities. *The Review of Economic Studies, 44*(3), 431-449.

Hartzmark, M. L. (1991). Luck versus forecast ability: Determinants of trader performance in futures markets. *The Journal of Business (Chicago, Ill.), 64*(1), 49-74. doi:10.1086/296525

Hendrickson, M. K., & James, H. S. (2016). Power, fairness and constrained choice in agricultural markets: A synthesizing framework. Journal of Agricultural & Environmental Ethics, 29(6), 945-967. https://doi.org/10.1007/s10806-016-9641-8

Hieronymus, T. A. (1977). Economics of futures trading for commercial and personal profit.

Hoffmann, A. O., Post, T., & Pennings, J. M. (2015a). How investor perceptions drive actual trading and risk-taking behavior. *Journal of Behavioral Finance*, *16*(1), 94-103.

Hoffmann, A. O., Post, T., & Pennings, J. M. (2015b). How investor perceptions drive actual trading and risk-taking behavior. *Journal of Behavioral Finance*, *16*(1), 94-103.

Kent State University Libraries. (2021). Independent samples t test.

Koziol, C., & Treuter, T. (2019). How do speculators in agricultural commodity markets impact production decisions and commodity prices? A theoretical analysis. European Financial Management : The Journal of the European Financial Management Association, 25(3), 718-743. https://doi.org/10.1111/eufm.12201

Leuthold, R. M., Garcia, P., & Lu, R. (1994a). The returns and forecasting ability of large traders in the frozen pork bellies futures market. *The Journal of Business (Chicago, Ill.), 67*(3), 459-473. doi:10.1086/296641

Leuthold, R. M., Garcia, P., & Lu, R. (1994b). The returns and forecasting ability of large traders in the frozen pork bellies futures market. *The Journal of Business (Chicago, Ill.), 67*(3), 459-473. doi:10.1086/296641

Lorton, S., & White, D. (2002). *Merchant's edge*. Champaign, IL: Stripes Publishing.

McKenzie, A. M., Isbell, B. J., & Brorsen, B. W. (2019). The cost of forward contracting in the CIF NOLA export bid market. *Journal of Agricultural and Applied Economics*, *51*(1), 164-181.

Microsoft.Microsoft excel [computer software]

Nicolosi, G., Peng, L., & Zhu, N. (2009). Do individual investors learn from their trading experience? *Journal of Financial Markets, 12*(2), 317-336.

Nofsinger, J. R. (2017). *The psychology of investing* (6th ed.). Milton: Routledge Ltd - M.U.A. doi:10.4324/9781315230856 Retrieved from https://www.taylorfrancis.com/books/9781351863315

Perloff, J. M., & Rausser, G. C. (1983). The effect of asymmetrically held information and market power in agricultural markets. *American Journal of Agricultural Economics, 65*(2), 366-372.

SAS Insititue Inc. (2019). Sas [computer software]. Cary, NC, USA:

SAS Institute.The FREQ procedure. Retrieved from https://documentation.sas.com/?cdcId=pgmsascdc&cdcVersion=9.4_3.5&docsetId=procstat&docsetTarget=procstat_freq_details08.htm&locale=en

Suhonen, N., & Saastamoinen, J. (2017). How do prior gains and losses affect subsequent risk taking? new evidence from individual-level horse race bets. *Management Science, 64*(6), 2797-2808. doi:10.1287/mnsc.2016.2679

Thaler, R. H., & Johnson, E. J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Science, 36*(6), 643-660.

UCLA: Statistical Consulting Group.Proc ttest | sas annotated output. Retrieved from https://stats.idre.ucla.edu/other/mult-pkg/faq/general/faq-how-do-i-cite-web-pages-and-programs-from-the-ucla-statistical-consulting-group/

Working, H. (1949). The theory of price of storage. *The American Economic Review, 39*(6), 1254-1262.

Tables

Table 1

Mean TM per trader over the whole dataset

Trader	Mean TM	Std Dev	n
11	0.27	0.47	3745
1	0.27	0.53	1638
10	0.26	0.18	375
2	0.26	0.53	933
4	0.24	0.53	2391
12	0.23	0.38	1803
13	0.22	0.21	546
9	0.18	0.17	4
7	0.18	0.27	282
5	0.16	0.52	9537
8	0.13	0.7	7000
3	0.11	0.15	18
6	0.03	0.19	1266

Table 2

Mean TM per Trader per year

Trader	Mean TM	Std Dev	n
12	0.36	0.44	4
2	0.3	0.28	4
4	0.29	0.29	4
11	0.26	0.1	6
10	0.26	0	2
1	0.23	0.18	7
9	0.18	0.2	2
3	0.18	0.13	2
7	0.15	0.1	2
13	0.14	0.1	4
8	0.12	0.14	4
5	0.12	0.11	8
6	-0.01	0.06	2

Table 3

Mean TM per Trader per quarter

Trader	Mean TM	Std Dev	n
2	0.36	0.48	12
4	0.33	0.29	11
12	0.26	0.34	8
11	0.25	0.13	21
1	0.24	0.25	23
9	0.18	0.2	2
10	0.18	0.15	4
7	0.16	0.03	2
13	0.16	0.08	7
8	0.14	0.15	12
5	0.14	0.15	26
3	0.13	0.16	3
6	0	0.05	5

Table 4

Detailed mean TM per trader per quarter

		Mean TM per Quarter (delivery)																									
Trader		Q01	Q02	Q03	Q04	Q05	Q06	Q07	Q08	Q09	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25	Q26
	Mean												0.4	0.1	0	0.1	0.2	0.1	0.2	0.2	0.2	0.3	0	-0			
	Std												0.8	2	0.7	0.4	0.3	0.4	0.4	0.4	0.4	0.5	0.4	0.3			
eight	N												40	270	329	336	357	344	357	357	337	280	302	191			
	Min												-1	-5	-2	-2	-1	-1	-1	-0	-0	-1	-1	-1			
	Max												2.6	3.6	2	1.2	0.9	1.7	1.3	1.3	1.4	2.6	1.2	1.1			
	Mean						0.1	0.2	0.4	0.5	0.5	0.4	0.2	0.5	0.1	0.2	0.1	0.3	0.3	0.1	0.1	0.2	0.2	0.2	0.2	0.3	0.3
	Std						0.2	0.2	0.8	0.6	0.5	0.4	0.6	1	0.3	0.3	0.3	0.3	0.3	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
eleven	N						3	8	34	104	148	159	232	160	183	199	228	208	208	181	192	200	198	198	147	97	85
	Min						-0	0	-1	-2	-1	-1	-2	-1	-1	-0	-2	-0	-0	-0	-0	-0	-0	-0	-0	-0	-0
	Max						0.3	0.6	2.3	1.7	2.5	2	2.3	3.8	0.7	1	0.6	1	1	1.2	0.7	0.8	0.8	0.7	0.7	0.7	0.7
	Mean	0	-0	-0	0.2	0.1	0	0.1	0.6	0.5	0.3	0.1	0.2	0.3	0.1	0.1	0.2	0.1	0.1	0.1	0	0.1	0.1	0.1	0.1	0.1	0.1
	Std	0.3	0.4	0.4	0.4	0.4	0.3	0.4	0.9	0.8	0.5	0.4	0.6	0.9	0.4	0.3	0.3	0.3	0.4	0.4	0.4	0.3	0.2	0.2	0.1	0.1	0.1
five	N	38	161	189	180	219	283	322	298	427	424	454	365	392	356	359	347	411	367	310	248	298	306	355	292	131	107
	Min	-1	-1	-1	-1	-2	-1	-1	-2	-2	-1	-1	-1	-6	-2	-1	-1	-1	-1	-1	-1	-1	-0	-0	-0	-0	-0
	Max	0.5	0.6	0.5	1.3	1.1	0.9	1.1	2.9	2.6	1.3	1.2	3.3	4.6	1.9	0.6	0.7	0.8	1.2	1.1	0.6	0.7	0.7	0.6	0.7	0.7	0.9
	Mean		0.3	0.3	0.4	-0	-0	-0	0.7	0.6	0.6	0.7	0.4														
	Std		0.4	0.4	0.5	0.5	0.2	0.4	0.7	0.6	0.7	0.5	0.1														
four	N		147	198	180	154	145	135	111	55	51	14	6														
	Min		-1	-1	-1	-1	-1	-1	-1	-1	-1	0.2	0.2														
	Max		1.3	1.2	1.4	1.1	0.5	1.1	2.6	2	1.8	1.8	0.5														
	Mean																									0	0.3
	Std																										0.1
nine	N																									1	2
	Min																									0	0.3
	Max																									0	0.4

Table 4 (cont.)

Mean TM per Quarter (delivery)

Trader		Q01	Q02	Q03	Q04	Q05	Q06	Q07	Q08	Q09	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25	Q26
	Mean	0.1	-0	-0	0.1	0.3	0.2	0.3	0.9	0.7	0.2	0.1	0.2	0.2	0.2	0.4	0.1		0.3	0.5		0.5	0	0.2		-0	0.3
	Std	0.3	0.3	0.3	0.4	0.3	0.3	0.3	0.8	0.9	0.2	0.3	0.2	0.3	0.2	0.2	0.1		0				0				
one	N	104	133	138	126	140	109	122	143	181	161	115	59	58	20	5	2		2	1		1	2	1		1	1
	Min	-1	-1	-1	-1	-1	-0	-0	-0	-1	-1	-1	-0	-0	-0	0.2	0		0.3	0.5		0.5	0	0.2		-0	0.3
	Max	0.6	0.5	0.7	0.8	0.9	0.9	0.9	3.2	2.8	1.1	1	0.8	0.9	0.6	0.7	0.2		0.3	0.5		0.5	0.1	0.2		-0	0.3
	Mean																									0.1	0.2
	Std																									0.1	0.3
seven	N																									15	126
	Min																									-0	-1
	Max																									0.3	0.8
	Mean																						-0	0.1	0	0	-0
	Std																						0.2	0.2	0.2	0.2	0.2
six	N																						9	127	210	219	68
	Min																						-0	-1	-1	-1	-1
	Max																						0.3	0.5	0.5	0.6	0.3
	Mean																							-0	0.3	0.3	0.3
	Std																							0	0.2	0.2	0.2
ten	N																							2	65	179	129
	Min																							-0	-0	-0	-0
	Max																							-0	0.7	0.8	0.7
	Mean																0.1					0.1	0.1	0.1	0.2	0.2	0.2
	Std																						0.2	0.2	0.2	0.2	0.2
thirteen	N																1					1	24	72	111	157	179
	Min																0.1					0.1	-0	-1	-0	-0	-0
	Max																0.1					0.1	0.6	0.6	0.7	1	0.7

Table 4 (cont.)

Trader		Q01	Q02	Q03	Q04	Q05	Q06	Q07	Q08	Q09	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25	Q26
		Mean TM per Quarter (delivery)																									
three	Mean						0.3					0.2	-0														
	Std											0.1	0.1														
	N						1					5	3														
	Min						0.3					0.1	-0														
	Max						0.3					0.4	0														
twelve	Mean																1				0	0	-0	0.4	0.4	0.2	0.1
	Std																				0.2	0.2	0.2	0.5	0.5	0.3	0.2
	N																1				19	54	48	147	228	226	179
	Min																1				-1	-0	-1	-1	-1	-1	-0
	Max																1				0.4	0.5	0.2	1.5	1.5	1.5	0.5
two	Mean	0.2	0	0	0.2	0.4	0.5	0.9	1.5	0.6	0.2	-0	-0														
	Std	0.4	0.5	0.5	0.5	0.4	0.5	0.6	0.6	0.5	0.1	0.3	0.4														
	N	107	160	162	159	154	79	38	10	19	32	10	2														
	Min	-1	-1	-1	-1	-1	-0	-0	0.2	0.1	-0	-1	-1														
	Max	0.8	0.9	1	1.2	1.1	1.1	1.4	2.3	1.4	0.4	0.2	-0														

Table 5

Winner and loser rankings for quarterly data based on average TM

t	t+1		winner t+1	loser t+1	N
Q02	Q03	winner t	2	0	4
		loser t	0	2	
Q03	Q04	winner t	2	0	4
		loser t	0	2	
Q04	Q05	winner t	1	1	4
		loser t	1	1	
Q05	Q06	winner t	2	0	4
		loser t	0	2	
Q06	Q07	winner t	2	0	5
		loser t	0	2	
Q07	Q08	winner t	2	0	5
		loser t	0	1	
Q08	Q09	winner t	2	0	5
		loser t	0	2	
Q09	Q10	winner t	0	2	5
		loser t	1	0	
Q10	Q11	winner t	2	0	5
		loser t	0	2	
Q11	Q12	winner t	1	0	6
		loser t	1	1	
Q12	Q13	winner t	0	2	4
		loser t	2	0	
Q13	Q14	winner t	1	1	4
		loser t	1	1	
Q14	Q15	winner t	1	1	4
		loser t	1	1	
Q15	Q16	winner t	0	2	4
		loser t	2	0	
Q18	Q19	winner t	1	1	4
		loser t	1	1	
Q20	Q21	winner t	2	0	4
		loser t	0	2	
Q21	Q22	winner t	0	1	6
		loser t	0	1	
Q22	Q23	winner t	0	0	7
		loser t	1	1	
Q23	Q24	winner t	1	0	6
		loser t	1	1	
Q24	Q25	winner t	1	1	5
		loser t	1	1	

Table 6

Pooled winner and loser rank test frequency table for quarterly data Based on average TM per Trader

Frequency Percent Row Pct Col Pct	Winner t+1	Loser t+1	Total
Winner t	23 31.94 65.71 63.89	12 16.67 34.29 33.33	35 48.61
Loser t	13 18.06 35.14 36.11	24 33.33 64.86 66.67	37 51.39
Total	36 50	36 50	72 100

Table 7

Results of the Fisher's exact test results for pooled quarterly data

Cell (1,1) Frequency (F)	23
Left-sided Pr <= F	0.9978
Right-sided Pr >= F	0.0089
Table Probability (P)	0.0067
Two-sided Pr <= P	0.0178

Table 8

Average TMs of the top and bottom traders for the quarterly data

	Average TM t+1 in USD per bu	
Quarter	Top-ranked trader in t	Bottom-ranked trader in t
Q2	0.26	-0.09
Q3	0.39	0.17
Q4	-0.1	0.32
Q5	0.53	-0.01
Q6	0.88	-0.02
Q7	1.5	0.73
Q8	0.58	0.47
Q9	0.17	0.52
Q10	0.69	0.06
Q11	0.36	-0.29
Q12	0.11	0.49
Q13	0.08	0.04
Q14	0.4	0.06
Q15	0.11	0.23
Q16	0.02	0.06
Q18	0.46	0.09
Q19	0.51	0.03
Q20	0.32	0.04
Q21	0.04	-0.09
Q22	0.16	0.35
Q23	0.38	0.25
Q24	0.22	0.04
Average	0.37	0.16

	Average TM t+1 in USD per bu	
Quarter	Average top two traders in t	Average bottom two traders in t
Q2	0.14	-0.07
Q3	0.3	0.2
Q4	0.16	0.25
Q5	0.36	0.02
Q6	0.52	0.06
Q7	1.21	0.93
Q8	0.66	0.73
Q9	0.2	0.55
Q10	0.53	-0.02
Q11	0.27	0.09
Q12	0.14	0.57
Q13	0.1	0.24
Q14	0.27	0.19
Q15	0.11	0.33
Q16	0.08	0.32
Q18	0.3	0.25
Q19	0.33	0.13
Q20	0.26	0.17
Q21	0.04	0.08
Q22	0.13	0.23
Q23	0.1	0.16
Q24	0.24	0.09
Average	0.29	0.25

Table 9

Top and bottom rank test of grain traders for the quarterly data

USD per bu	Average TM t+1	t-statistic p-value	Wilxocon p-value
Top trader	0.37		
Bottom trader	0.16		
Top 2 traders average	0.29		
Bottom 2 traders average	0.25		
Top versus Bottom	0.21	0.0146	0.0182
Top 2 versus Bottom 2	0.04	0.0529	0.125

Table 10

Descriptive Statistics of average TM per trader by year

	Mean TM per Year (delivery)																			
	2011					2012					2013					2014				
Trader	Mean	Std	N	Min	Max	Mean	Std	N	Min	Max	Mean	Std	N	Min	Max	Mean	Std	N	Min	Max
eight																0.25	1.6	864	-5.3	3.6
eleven											0.42	0.6	351	-2.3	2.3	0.41	0.7	1102	-2.4	3.8
five	-0	0.2	58	-0.6	0.5	0.03	0.4	1414	-1.7	1.3	0.24	0.7	2058	-2.1	2.9	0.23	0.7	2141	-5.9	4.6
four	0		1	0	0	0.2	0.5	1450	-1.3	1.4	0.26	0.6	842	-1.1	2.6	0.69	0.6	98	-0.3	1.8
nine																				
one	0.07	0.3	140	-0.8	0.6	0.1	0.4	538	-1.1	0.9	0.56	0.7	573	-1.1	3.2	0.14	0.3	361	-0.9	1.1
seven																				
six																				
ten																				
thirteen																				
three						0.27	0	2	0.27	0.3						0.09	0.2	16	-0.1	0.4
twelve																				
two	0.21	0.4	150	-0.9	0.8	0.19	0.5	626	-1	1.2	0.72	0.6	123	-0.4	2.3	0.1	0.2	34	-0.7	0.3

Table 10 (cont.)

	Mean TM per Year (delivery)																			
	2015					2016					2017					2018				
Trader	Mean	Std	N	Min	Max	Mean	Std	N	Min	Max	Mean	Std	N	Min	Max	Mean	Std	N	Min	Max
eight	0.08	0.4	2760	-2.1	1.7	0.2	0.4	2578	-1.1	2.6	-0.1	0.4	798	-1	1.2					
eleven	0.17	0.3	857	-1.7	1	0.18	0.3	770	-0.4	1.2	0.2	0.2	598	-0.5	0.7	0.24	0.2	67	-0.3	0.6
five	0.15	0.3	1579	-1.9	1.9	0.09	0.3	1187	-1	1.1	0.11	0.2	1042	-0.3	0.7	0.1	0.1	58	-0.1	0.9
four																				
nine											0.03	0	2	0.03	0	0.32	0.1	2	0.26	0.4
one	0.31	0.2	17	-0.1	0.7	0.31	0.2	5	0.03	0.5	0.09	0.2	4	-0.2	0.3					
seven											0.08	0.1	74	-0.2	0.3	0.22	0.3	208	-1	0.8
six											0.03	0.2	1210	-0.7	0.6	-0.1	0.2	56	-0.3	0.3
ten											0.26	0.2	301	-0.1	0.8	0.26	0.2	74	-0.3	0.7
thirteen	0.05	0	2	0.05	0.1	0.06	0.1	8	0	0.2	0.22	0.2	409	-0.5	1	0.23	0.2	127	-0.2	0.6
three																				
twelve	1	0	2	1	1	0.03	0.2	192	-0.7	0.5	0.28	0.4	1361	-1	1.5	0.12	0.2	248	-0.3	0.5
two																				

Table 11

Winner and loser rankings for yearly data based on average TM

t	t+1		winner t+1	loser t+1	N
2012	2013	winner t	1	1	4
		loser t	1	1	
2013	2014	winner t	0	2	5
		loser t	1	0	
2014	2015	winner t	1	1	4
		loser t	1	1	
2015	2016	winner t	1	1	6
		loser t	1	1	
2016	2017	winner t	0	2	6
		loser t	2	0	

Table 12

Pooled Winner and Loser Rank Test Frequency Table for Yearly Data Based on Average TM per Trader

Frequency Percent Row Pct Col Pct	Winner t+1	Loser t+1	Total
Winner t	3 15.79 30 33.33	7 36.84 70 70	10 52.63
Loser t	6 31.58 66.67 66.67	3 15.79 33.33 30	9 47.37
Total	9 47.37	10 52.63	19 100

Table 13

Results of the Fisher's exact test for pooled yearly data

Cell (1,1) Frequency (F)	3
Left-sided Pr <= F	0.1276
Right-sided Pr >= F	0.9815
Table Probability (P)	0.1091
Two-sided Pr <= P	0.1789

Table 14

Average TMs of the top and bottom traders for the yearly data

	Average TM t+1 in USD per bu			Average TM t+1 in in USD per bu	
Year	Top-ranked trader in t	Bottom-ranked traders in t	Year	Average top two traders in t	Average bottom two traders in t
2012	0.09	0.33	2012	0.17	0.45
2013	0.10	0.69	2013	0.12	0.45
2015	0.17	0.31	2015	0.13	0.23
2016	0.03	0.06	2016	0.17	0.13
2017	0.09	0.28	2017	0.01	0.25
Average	0.09	0.33	Average	0.12	0.30

Table 15

Top and bottom rank test of grain traders for the quarterly data

USD per bu	Average TM t+1	t-statistic p-value	Wilxocon p-value
Top trader	0.09		
Bottom trader	0.33		
Top 2 trader's average	0.12		
Bottom 2 trader's average	0.3		
Top versus Bottom	-0.24	0.0649	0.0625
Top 2 versus Bottom 2	-0.18	0.0529	0.125

Table 16

Average TM for 2017 by gender and trader

Gender	Trader	TM (USD per bu)				
		Mean	Std	N	Min	Max
female	six	0.03	0.19	1210	-0.72	0.62
	thirteen	0.22	0.21	409	-0.53	1.04
	twelve	0.28	0.41	1361	-0.98	1.52
male	eight	-0.07	0.39	798	-1.04	1.21
	eleven	0.2	0.2	598	-0.45	0.72
	five	0.11	0.15	1042	-0.25	0.7
	nine	0.03	0	2	0.03	0.03
	one	0.09	0.2	4	-0.17	0.3
	seven	0.08	0.14	74	-0.18	0.34
	ten	0.26	0.18	301	-0.12	0.79

Table 17

T-test of Trader Gender Performance for 2017 Based on Average TM per Bushel

Variable: TM in USD per bu

Gender	N	Mean	Std Dev	Std Err	Minimum	Maximum
Female	2980	0.1697	0.3298	0.00604	-0.9812	1.5233
Male	2819	0.0945	0.2802	0.00528	-1.0437	1.2074
Diff (1-2)		0.0752	0.3067	0.00806		

Gender	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
Female		0.1697	0.1579	0.1816	0.3298	0.3216	0.3384
Male		0.0945	0.0842	0.1049	0.2802	0.2731	0.2877
Diff (1-2)	Pooled	0.0752	0.0594	0.091	0.3067	0.3012	0.3124
Diff (1-2)	Satterthwaite	0.0752	0.0595	0.0909			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	5797	9.33	<.0001
Satterthwaite	Unequal	5731.9	9.38	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	2979	2818	1.39	<.0001

Table 18

Average TM by gender and trader for the whole dataset

Gender	Trader	TM (USD per bu)				
		Mean	Std	N	Min	Max
Female	six	0.03	0.19	1266	-0.72	0.62
	thirteen	0.22	0.21	546	-0.53	1.04
	three	0.11	0.15	18	-0.11	0.4
	twelve	0.23	0.38	1803	-0.98	1.52
Male	eight	0.13	0.7	7000	-5.26	3.61
	eleven	0.27	0.47	3745	-2.36	3.75
	five	0.16	0.52	9537	-5.94	4.56
	four	0.24	0.53	2391	-1.32	2.58
	nine	0.18	0.17	4	0.03	0.38
	one	0.27	0.53	1638	-1.11	3.19
	seven	0.18	0.27	282	-0.99	0.76
	ten	0.26	0.18	375	-0.27	0.79
	two	0.26	0.53	933	-1.03	2.26

Table 19

T-test of Trader gender performance based on average TM over the whole dataset

Variable: TM in USD per bu

Gender	N	Mean	Std Dev	Std Err	Minimum	Maximum
Female	3633	0.1574	0.3128	0.00519	-0.9812	1.5233
Male	25556	0.1879	0.5682	0.00355	-5.9438	4.5616
Diff (1-2)		-0.0305	0.543	0.00963		

Gender	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
Female		0.1574	0.1472	0.1676	0.3128	0.3058	0.3202
Male		0.1879	0.181	0.1949	0.5682	0.5633	0.5732
Diff (1-2)	Pooled	-0.0305	-0.0494	-0.0117	0.543	0.5387	0.5475
Diff (1-2)	Satterthwaite	-0.0305	-0.0429	-0.0182			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	29187	-3.17	0.0015
Satterthwaite	Unequal	7601.4	-4.85	<.0001

Equality of Variances

Method	Num DF	Den DF	F Value	Pr > F
Folded F	25555	3632	3.3	<.0001

Table 20

Average TM for 2017 by experience and trader

Experience	Trader	TM (USD per bu)				
		Mean	Std	N	Min	Max
Three years or less	nine	0.03	0	2	0.03	0.03
	seven	0.08	0.14	74	-0.18	0.34
	six	0.03	0.19	1210	-0.72	0.62
	ten	0.26	0.18	301	-0.12	0.79
	thirteen	0.22	0.21	409	-0.53	1.04
	twelve	0.28	0.41	1361	-0.98	1.52
Four years or more	eight	-0.07	0.39	798	-1.04	1.21
	eleven	0.2	0.2	598	-0.45	0.72
	five	0.11	0.15	1042	-0.25	0.7
	one	0.09	0.2	4	-0.17	0.3

Table 21

T-test of Trader experience for 2017 based on average TM

Variable: TM in USD per bu

Trader Experience	N	Mean	Std Dev	Std Err	Minimum	Maximum
3 years or less	3357	0.1757	0.3175	0.00548	-0.9812	1.5233
4 years or more	2442	0.0747	0.2868	0.0058	-1.0437	1.2074
Diff (1-2)		0.1009	0.305	0.00811		

TraderExperience	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
3 years or less		0.1757	0.1649	0.1864	0.3175	0.3101	0.3253
4 years or more		0.0747	0.0634	0.0861	0.2868	0.279	0.2951
Diff (1-2)	Pooled	0.1009	0.085	0.1168	0.305	0.2995	0.3106
Diff (1-2)	Satterthwaite	0.1009	0.0853	0.1166			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	5797	12.44	<.0001
Satterthwaite	Unequal	5534	12.64	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	3356	2441	1.23	<.0001

Do Reference Points Impact Forward Trading Behavior?

Abstract

A Two-way panel data model is applied to test the use of reference points of commercial agricultural commodity traders in their purchasing decisions. We test if either the total amount of bushels purchased by a specific trader for a specific feed mill location or the length of forward contracts are influenced by either the basis or the previous performance of the specific trader. We use monthly and quarterly time horizons. Our study differs from most other studies in the field of trading that focuses on either institutional investors or private speculators by investigating the behavior of commercial commodity traders in a risk-management setting. The traders in this study do not use their own wealth when engaging in a trade but are responsible for the steady input supply to avoid interrupting the feed supply chain for meat manufacturing companies. Results from our analysis indicate that traders do not change the amount of bushels purchased based on basis levels for their market locations, nor based on their trading performance in the previous period. However, we find that traders are influenced by the basis levels for their market locations when deciding on the contract length of forward purchases.

Keywords: futures markets; forward contracts; commodity purchasing

Introduction

Findings in behavioral finance have discovered that if many traders suffer from the same psychological bias, it can affect markets (Coakley, Dollery, & Kellard, 2011; Coval & Shumway, 2005; Nofsinger, 2018). Such findings highlight the increasing importance of studying the conscious and unconscious decision-making progress of all market participants.

Capitalist societies depend on functioning markets for the trade of goods and services. A unique role is often assigned to agricultural markets to guarantee food security. Food security is also why agriculture is a primary focus of policymakers who often allocate large budgets for the functioning of a country's agricultural sector. Until the second half of the twentieth century, the efficient market hypothesis (EMH) by Eugene Fama (1970) and the expected utility theory were the dominating theories to study the behavior of market participants. Fama defined a market as efficient when prices reflect all available information entirely at all times. Under the EMH, no market participant can systematically beat the market. Later in the 2nd half of the 20th century, the EMH and its underlying assumptions were challenged, leading to behavioral economics' emergence (Shleifer, 2000).

The publication of prospect theory by Kahneman and Tversky in 1979 has been a significant milestone for the field of behavioral economics and raised continued interest in the understanding of human behavior in financial markets. Kahneman & Tversky (1979) described that prospect theory focuses on reference points and problem representation to code outcomes of gambles as losses or gains. Further, prospect theory showed how humans value gains and losses differently, as previously assumed by the EMH. Ultimately, the value function of prospect theory can be described as being S-shaped, concave in shape for gains, and convex for losses, with the loss side being steeper than the gains side. The shape of the value functions leads to asymmetric decision-making based on losses looming larger than gains (Kahneman & Tversky, 1979). Two biases resulting from prospect theory's value function are loss aversion (Thaler, Tversky, Kahneman, & Schwartz, 1997) and the disposition effect (Shefrin & Statman, 1985).

While many discoveries have been made so far about behavior and the decision-making process in the financial field in general, findings in the field of commodity trading are sparser. Commodity markets, especially for soft commodities or agricultural products, follow rules that do not necessarily apply to other markets. For example, agricultural products are usually seasonal and perishable, which is unique in financial markets, especially compared to stocks. Additionally, commodity contracts on futures exchanges are traded with contracts that specify and standardize the trades. All transactions on futures exchanges have a buy or sell obligation for an exchange at a date in the future. Traditionally, futures market transactions for commodities serve a risk management purpose in the form of hedging. Moreover, similarly, forward contracts for the physical exchange of commodities help producers and purchasers lock in a price to sell or buy the product. An agreed-upon price helps both sides, buyers, and sellers achieve planning certainty and reduce financial risk.

Background

This research focuses on the behavior of commodity traders employed by a firm producing livestock feed for the manufacturing of meat products. In this case, only corn transactions are considered as an input factor to the feed products. Profit margins in meat production are typically tight (Martin, Smith & Smith, 2016), making profitable feed input purchases critical to the firms' overall profitability. To manage the price risk, the traders in the focus of this study use different contracts to lock in purchase prices of corn for future deliveries.

Those three contract types are futures contracts, forward contracts, and basis contracts. The trading team of the firm consists of two groups. The first group are traders that focus on transactions in the futures market. Futures market traders have a specific skill set in timing transactions in the futures market and purchasing futures positions with the target to mitigate price risk. Those traders timing the futures market purchases are not the focus of this study. The behavior of the grain traders in the second group, who are the focus of this study, have skills that may not be easily compared to traders in other markets, such as the financial market. For example, traders in commercial grain trading are using basis rather than the price to evaluate transactions.

Basis (B_t) represents the difference between the commodity's cash price (CP_t) and the respective futures price (FP_t):

$$B_t = CP_t - FP_t \tag{1}$$

The basis represents different costs to bring the commodity to a specific market at a particular time. For example, when the futures price for the May contract for corn is trading at $4.90 and the cash price in Purdy, MO is $4.50, the basis is -$0.40, the basis is 40 cents under the May futures. If the cash price were $5.00, the basis would be 10 cents over May in the same scenario. Fuel and transportation costs determine basis levels in addition to financial and physical and storage costs.

Our data uses a basis adjusted for transportation costs, called derived gateway basis, to make transactions comparable regardless of their location. In our case, to mitigate the upward price risk of CP_{t+1} for location *j*, basis contracts are used to create hedged positions that remove all conditional price risk. When the hedged positions are established, the traders in our focus engage in basis contracts that, in combination with the futures contracts, set a cash price. If the trader can secure a low basis, they are able to set a low effective hedged futures price. A more detailed description of the basis trading background can be found in essay one.

In this study, we focus on the past performance of the traders, which can be measured by trader margin *(TM):*

$$TM = B_{t+1} - B_t^{t+1} \tag{2}$$

Where *TM* represents the difference between the actual basis at physical delivery time *t+1*, denoted as B_{t+1}, and the basis established through a basis contract in period *t*, denoted as B_t^{t+1}. *TM* provides information on how good or bad a forward trade was compared to a purchase in the spot market at the time of delivery based on the basis of a transaction at delivery time t+1 and the basis contract at time t. For example, a negative *TM* would indicate that the basis the trader locked in in period t with a basis contract plus was higher than the basis at delivery. Like basis, the *TM* is comparable by trader and for different market locations.

The grain traders that are the focus of this study are tasked with deciding when to basis contract grain and for how much grain to use basis contracts at specific times throughout a seven-year sample period of August 2011 to February 2018. In our study traders purchase grain for different locations which is accounted for by analyzing each trader and location combined as a separate cross-section.

Our study will seek to determine if decision-making under uncertainty by individual traders employed by the firm is influenced by various heuristics drawn from behavioral economics. Although basis contracts are designed to remove purchase price risk for a firm buying input products like corn, the timing of when and how much to forward contract basis is a subjective decision taken by individual traders. Furthermore, given that the level of an individual trader's *TM* is determined by this timing decision, from an individual trader's perspective, the length and volume of grain in a forward contract is related to the trader's expected *TM*. In this sense, the volume and contract length of forward contracts contain inherent risk for the trader, and the decision of when and how much to basis contract may be influenced by behavioral biases linked to past performance, as well as to current and future expected market conditions. If either the size or the range of the contract increases, so does the attached risk. The forward trades' performance is measured by the *TM* that is achieved after the transaction took place. Since the contracts are for delivery in the future, prices can fluctuate until the final delivery. Thus, the trader might purchase grains for future delivery at a higher cost than the spot market price at delivery. If a trader paid a higher price for corn with a forward purchase than the spot market price at delivery, the *TM* would be negative. Because a long-time span for a forward delivery provides more time for the price to fluctuate, a more extended contract may be riskier. It can be assumed that if a trader agrees on a contract for delivery in the future, the larger the volume and the longer the contract length, the more comfortable the trader is that the contract agreement is a good deal for them. This study will help gain new insights into grain merchandising behavior to support government agencies and consulting companies in creating new marketing strategies and improving existing ones.

Literature

The question arises as to how humans behave in uncertain situations. Kahneman and Tversky's famous prospect theory (1979) offers a suitable framework to understand behavior in risky choice situations better. Prospect theory pushed research from utility theory and wealth status of individuals to investigating how humans feel about a change in wealth, as was previously suggested by Markowitz (Kahneman, 2011, page 278). Three significant discoveries of studying the changes of wealth were made by Kahneman and Tversky (1979). The first discovery is the status quo or reference point. Humans use a reference point to frame their situation that ultimately guides their decisions to evaluate gains and losses. Mattos & Zinn (2016) studied reference points in agriculture, who are the first to explore reference points in commodity marketing. They used an experimental setting to test how grain producers use reference prices during their marketing decisions. Their findings show that current prices and price trends influence the farmer's decision-making, and also price expectations play a role in their decision process.

A German study about anchoring effects among farmers with experimental auctions was carried out in 2013. It was able to find that in a series of auctions with the same group of farmers, the previously successful bid influenced the following auction. The authors suggested that bids' adjustment based on the previous winning bid acts as an endogenous anchoring point. In this case, an update of the farmers' preferences is possible, but the update would happen regarding the previously successful bid, which would again be an anchor. Overzealous bidding was also not found in this study, where farmers would have bet higher to win the auctions rather than letting their group members have a chance of winning an auction (Holst, Herman, & Musshoff, 2015).

The second discovery concerns the diminishing sensitivity to a change in wealth which is the foundation for the third discovery, loss aversion. For example, to explain the diminishing sensitivity to a change in wealth, a gain of 100 United States Dollars (USD) feels more critical when the wealth changes from 100 USD to 200 USD compared to a 100 USD grain from 1,000 USD to 1,100 USD. The third prospect theory discovery is understanding how gains and losses are perceived, leading to loss aversion. Humans perceive the pain felt from a loss worse than the joy of an equal gain would provide. Ultimately the behavior can be observed when people decline to bet on a coin for equal stakes (Kahneman & Tversky, 1984). In a fair coin gamble with equal stakes, the possible loss makes the possible gain unattractive.

Loss aversion has further behavioral implications. In 1985, Shefrin, Hersh M. & Statman (1985) first described the disposition effect, who found that investors tend to hold on too long to losing positions hoping that they turn into winners based on an aversion to realizing losses. On the other hand, investors are too quick to sell their winning positions. The disposition effect was studied among wheat farmers by Mattos & Fryza (2014) to learn more about the marketing performance of wheat farmers. Their study results provide evidence of the disposition effect among wheat farmers, where farmers are more eager to sell if prices are above the reference price.

Beyond the disposition effect, loss aversion also plays a significant role in behavior spanning multiple periods. It has been shown that humans do not always make rational decisions, and decision-making is getting even more complicated when more than one time period has to be considered. Thus, the evaluation of a current situation varies and depends on the historic experiences the trader has made. A 2005 study by Coval & Shumway found evidence of loss aversion among Chicago Board of Trader proprietary traders. Their study was among the first studies to test actual rather than experimental behavior.

Prior grains from risky gambles may lead to the so-called house money effect when traders engage in more risky gambles after being initially successful. In contrast, previous losses can lead to a snake-bite effect (Nofsinger, 2017), also known as the playing-safe effect (Suhonen & Saastamoinen, 2017). Both effects occur when there is very little or no chance to recover previous losses with additional gambles. Previous losses can also lead to the trying-to-break-even effect (Nofsinger, 2017) when the situation offers a considerable chance to make up for prior losses. Thus, traders who have recently experienced losses try to either make up for their losses with more risky transactions or shy away from risk based on the expected outcomes of the consecutive offered gambles.

Consequently, the question arises of how prior losses and gains affect risk-taking behavior. Thaler and Johnson (1990) asked this question and found that "prior gains and losses can dramatically influence subsequent choices in systematic ways" (p. 643). The authors conducted a study with students and tested for evidence of different mental processing strategies called editing rules.

Framing plays a significant role in evaluating gains and losses. When presenting or processing information, the status quo or an initial wealth level must be set to evaluate possible outcomes. The status quo is vital to project outcomes of transactions as gains or losses (Kahneman & Tversky, 1984). Framing happens during the editing phase, which is an essential component of prospect theory. In the editing phase, individuals find their reference point and simplify and encode their chances. Prospect theory uses editing with and without memory.

Hedonic editing assumes that humans edit the presented situation to make the prospects appear most pleasant or unpleasant. Based on a previous study of Thaler (1980), the suggested principles of hedonic editing are 1. Segregate gains, 2. Integrate losses, 3. Segregate small gains from more considerable losses, and 4. Integrate more minor losses with more enormous gains. One primary observation they make is that segregation is facilitated when longer time-spaces separate two events. The students taking the survey indicated that they would prefer to spread out pleasant events to several days, probably to extend the pleasure that those events cause.

Following the same thoughts, the students should have preferred to bundle losses in one day, but they instead preferred to spread the hurtful events to separate days. Those findings indicate that losses are not integrated on the same day and that the second loss on a single day is more painful than isolated on a separate day. Thaler and Johnson scrutinized their results and found evidence that suggests that their subjects use a so-called quasi-hedonic editing rule, which builds on subsequent outcomes. Following prospect theory, for simple prospects, humans are risk-seeking when facing losses. This experiment revealed that the integration of losses is not always happening, especially for more complex prospects. Here, risk aversion was increased by 60% of the students when winning the second offered gamble would not offset the initial losses. The students shied away from the pain, and additional loss would possibly cause the shying away from further losses known as the snake-bite effect. Contrarily, suppose the proposed gamble offers a chance to make up for the initial losses. In that case, risk acceptance will increase if the proposed gamble offers a chance to make up for the initial losses. The break-even effect induces the integration of prior losses as humans mentally eliminate their losses.

Quasi-hedonic editing also shows proof for a so-called house money effect. Humans are risk-seeking and experience suppressed loss aversion when successive losses are smaller than a prior gain. In this case, losses are immediately integrated with the prior gain. Further, Battalio, Kagel, and Jiranyakul (1990) suggest that the house money effect decreases if the possible subsequent loss gets closer to the initial stake.

In a multi-period experiment with students, Ackert, Charupat, Church, and Deaves (2006) found that wealth changes are negligible while the absolute level of wealth is the dominating influence for bidding behavior on stocks in a subsequent period. The conclusion was drawn after the students were endowed with money before the experiment and showed bidding behavior consistent with the house money theory. Still, winning students were unwilling to bid higher prices for an asset after an increase in wealth. While observing the horse racetrack's betting behavior, Suhonen and Saastamoinen (2017) found that the gamblers experiencing losses earlier in the race day show reduced risk-taking behavior as the day progresses. The study also showed decreased risk-taking when the gamblers tried to break even, which is contrary to the common belief that gamblers take on risky bets to increase their chances of offsetting their previous losses. (Friesen, Jeutand, & Unlu, 2021) study decision making under uncertainty among CEOs while comparing past unsuccessful repurchasing decisions of stocks on future stock repurchasing decisions. They found strong evidence that losses from past repurchases influence repurchasing decisions. The snake-bite effect could explain this behavior; traders do not want to experience a bite repeatedly. The authors provide further insight into the effects of trader experience and compensation.

Jaba, Robu, & Balan (2017) conducted performance assessments in the financial field using panel data analysis. While their study focused on evaluating financial companies' performance based on the return on equity and financial leverage, their method can be used to evaluate individuals' performance.

Concerning our dataset, we can observe trader behavior based on recent past performance at any point in time. For example, if a trader made losses measured by his recorded *TM* in the last period, does that impact their decision to forward contract in the present period? A reference or status quo point has to be set to capture how traders evaluate their performance. For example, in the financial literature, this is usually the asset's purchase price, also used by Odean (1998). In this case, this measure is not applicable, which is why the margin per trader and period was chosen as a performance measure. A similar measure was used by Friesen et al. (2021), who used the accumulated gains and losses of stock repurchases over different periods as reference points for their study. Since we do not know how grain traders self-evaluate their performance, the quarterly periods are aligned with corn's natural growth and harvest cycle. Monthly periods are also tested to test if a shorter period is more appropriate to capture trader behavior. Some traders would have been recent winners and some recent losers. A panel data regression was consulted to test how past performance influences forwarding decisions regarding quantity and length of contracts. Results can be compared across traders during the same time interval.

Data

The original dataset for this analysis consists of single purchase transactions of soybean meal and corn of a private company producing livestock feed. Each transaction information includes which trader set up the trade, the dates of contract set up, expected delivery dates, actual delivery dates, margin, and many more variables that were not important for this research project, such as transportation mode. The dataset comprises all transactions between August of 2012 to February 2018.

In the first step, all soybean transactions were dropped from the dataset because of a lack of observations. In the second step of cleaning the data, the transactions made by an actual farmer were excluded. The decision to eliminate farmer-initiated trades was made because farmers determine the forward contracting decision-making (timing and volume). After examining the data, three corn traders had to be excluded because of the low trading activity and the trading periods not being consecutive.

Now the reduced dataset had to be further edited for data analysis. For example, the prices in the dataset were not all in the same unit. While prices are in USD per bushel, the deliveries' unit is in cents per pound. For uniformity, all units were set to bushels and USD per bushel.

Further, it is crucial to incorporate a time component into the data to measure forward contracts. To include the time component, the number of days between the contract set up and the delivery were calculated. After we determined the number of days between contract set up and delivery, all trades with seven or more days between contract set-up and delivery are labeled as forward contracts. Forward contracts are deals that have a delivery date in the future as part of a purchase contract and are a tool often used in commodity trading to reduce risk.

Since the panel data procedure requires strict data sorting, several steps were undertaken to fit the data further. First, additional variables were created that extracted the year and month from the delivery date and the unloading date to help with SAS's sorting requirements. Then all trades were grouped by the month in which deliveries took place, and in a second set, all trades were grouped by the month that contracts were set up. These steps were necessary to calculate past performance and current purchasing behavior. The same was done for quarterly periods. The quarterly periods were set per the production cycle of corn. The quarters are three months periods from 1st of September to 30th of November, 1st of December to 28th of February, 1st of March to 31st of May, and 1st of June to 31st of August.

The final cross-sections for the panel data regression were created by grouping trades by trader i and by designated delivery location j. 70 designated delivery locations are in the original dataset. Our analysis does not consider location effects. For most locations, a single trader is responsible for repeated transactions at a time.

Next, the average derived gateway basis was calculated for all forward contracts during a trader's purchasing month or quarter for a specific location. The derived gateway basis is the contracted basis adjusted for freight obtained by a trader at the contract set-up date. Usually, the basis is defined as the difference between the local spot price for a commodity and the futures price for the same commodity (Lorton & White, 2002). The basis differs locally based on market supply and demand situations and varies over time. For example, grain traders use the basis to conduct business instead of the grain price, as prices themselves can be very volatile. Derived gateway basis means the basis was adjusted for fright or backed off to a gateway and makes basis across different mill locations comparable. For example, if the basis was -10 Dec at a location (mill) where grain was purchased and the transportation cost was 10 cents/bu to the nearest gateway location, the actual derived gateway basis for the mill location would be reported as 0 Dec basis. The last calculation was the average *TM* a trader achieved based on all purchases in a delivery month. Since we are interested in the trader's performance in the past, this average *TM* has is lagged one time period (month or quarter), which is the average *TM* achieved by a trader in the month or quarter before a respective purchase. The new lagged average *TM* is called $AvgTMMonth_{t-1}$ or $AvgTMQuarter_{t-1}$.

Methods

A panel data analysis approach was used to explore information across traders and locations and time. Panel data consists of two dimensions. The first being a cross-section, *i*, and the second a time-series dimension, *t*. Typically individuals, countries, or companies for which several time-series observations are given (Hsiao, 2007). Ajmani (2009) describes in an example why panel data is used to analyze the performance of financial advisors in a financial planning company. Several variables for financial advisors may explain the variability in profits of the financial advisors. Vivek (2009) names variables that are observed and controllable, like the client's wealth, the age of the trader, or the geographical location. An example of an unobserved variable would be "salesmanship ability" (P. 110). This example matches the commercial grain traders in this research, which have similar characteristics. Still, instead of salesmanship ability, the grain traders are expected to have excellent forecasting ability regarding the movement of commodity basis. Another distinction between the explanatory variables can be drawn between time-dependent and time-independent. Gender, for example, would be categorized as time-independent, while the job experience would be time-dependent.

Complete panels have the same amount of observations for each cross-sections and time-sections, meaning that each cross-section has an observation in each time-period in the dataset. While this represents the ideal case, many panels in real-life situations have missing observations. Panels with missing observations are called unbalanced panels. In the present case, the panel is unbalanced. The unbalanced data is not an issue for the SAS software, which accounts for unbalanced data for all the applied methods. (SAS Institute, 2017). The software detects missing data for the time series identification or other variables and treats the observation as missing for the respective cross-section. SAS will ignore observations with missing values for the model fit.

Besides balanced and unbalanced, models can be separated into one-way or two-way models, where one-way models depend on the cross-section or the time-section. In contrast, two-way models rely on the cross-section and the time section of the observation (SAS Institute, 2017).

A standard panel data model for each *i* with unobserved effects follows equation 3 (Wooldridge, 2016):

$$y_{ijt} = \beta_0 + \beta_1 AvgDerivedGWBasisMonth_{ijt} \\ +\beta_2 AvgTMMonth_{ijt-1} + \alpha_{ijt} + u_{ijt} \tag{3}$$

where the dependent variable y_{it} is either the total amount of bushel purchased with forward contracts or the contract length with respect to trader *i and* location *j* in time period *t*. β_0 is a constant and β_1 and β_2 are the parameters to estimate, in our case, the derived gateway basis and the lagged average margin of a trader *i* for a specific location *j* each for a specific time period *t*. The unobserved effect is α_{ijt}, and the idiosyncratic, or individual, error u_{ijt}.

For a two-way model, the error component would look a little different with consideration of unobservable time effects, as can be seen in Equation 4 (Baltagi, 2008):

$$y_{ijt} = \beta_0 + \beta_1 AvgDerivedGWBasisMonth_{ijt} + \beta_2 AvgTMMonth\ _{ijt-1} + \alpha_{ijt} + u_{ijt} \\ where\ u_{ijt} = \ k_{ijt} + \lambda_{ijt} + \nu_{ijt} \tag{4}$$

Here the error component u_{ijt} consists of k_{ijt}, which is now the unobservable individual effect, λ_{ijt} denotes the unobserved time effect, which is individual invariant, and lastly ν_{ijt} is the remaining stochastic disturbance term.

Panel data models can control for unobserved heterogeneity with fixed-effects models. Fixed-effects explore the relationship between dependent and independent variables within an individual. The fixed-effects method allows for an arbitrary correlation between unobserved effects α_i and x_{ijt}, which random effect models do not allow. The fixed-effects enable to treat uncontrolled subject-specific effects to be treated as constants (Ajmani, 2009). While removing the impact of the personal characteristic, fixed-effects can assess the effect of the independent variables. It has to be noted that fixed-effects models cannot handle time-invariant independent variables, such as gender. Three different approaches exist for fixed-effects models estimation, which are (1) the least squares dummy variable model (LSDV), (2) the within-group-, and (3) the between-group effects approach.

Two tests can be conducted to determine if a fixed-effects model was suitable for the data. First, included with the output generated by SAS for the panel data analysis is the F-test. For panel data models, The F-test tests for the poolability across cross-sections. The null hypothesis states that all fixed effects are jointly zero. For the F-test, the pooled model functions as a restricted model with the restriction that fixed heterogeneity terms must be constant for all cross-sections. Second, the F-test compares the fixed-effects estimates to pooled regression estimates (Ajmani, 2009; Greene, 2003). The formula for a one-way F-test can be found in equation 5:

$$F = \frac{(SSE_r - SSE_u)/df_1}{\frac{SSE_u}{df_2}} \sim F(df_1, df_2) \tag{5}$$

SSE_r is the error sum of squares from the restricted pooled estimation, and SSE_u represents the error from the unrestricted LSDV estimation. The degrees of freedom (df) in the case of one-way fixed effects have a numerator of N-1, while the numerator for two-way fixed-effects is (N-1)+(T-1) (SAS Institute, 2017). Therefore, if the test result is lower than alpha, then a fixed-effects model is the preferred estimation method.

Further, a Hausman specification test (Hausman, 1978) was invoked. This test is also known as m-statistic. The Hausman specification test tests if idiosyncratic errors are correlated with the independent variables. If the test result is significant, we reject the null hypothesis of no correlation, and fixed-effects are considered the suitable model. The Hausman specification is calculated in Equation 6, where S_2 and S_1 are consistent estimates of the asymptotic covariance matrices of β_1 and β_2. In this case, m is distributed $\chi 2$ with k degrees of freedom, where K is the dimension of β_1 and β_2.

$$m = (\beta_2 - \beta_1)'^{(S_1 - S_2)^{-1}}(\beta_2 - \beta_1) \tag{6}$$

Under the H_0, both estimators β_1 and β_2 are consistent. However, under the H_A, only β_2 is consistent, which leads to declining the use of random effects (SAS Institute, 2017) P.1873.

The Breusch and Pagan (1980) Lagrange Multiplier (LM) test can be used to test for cross-sectional dependence. The SAS software can easily access this test and test the null hypothesis that the data contains zero cross-sectional error correlations. The SAS user manual (SAS Institute, 2017) describes the test as:

$$\hat{p}_{ij} = \hat{p}_{ij} = \frac{\sum_{t=\underline{T}_{ij}}^{\overline{T}_{ij}} e_{it}e_{jt}}{\sqrt{\sum_{t=\underline{T}_{ij}}^{\overline{T}_{ij}} e_{it}^2}\sqrt{\sum_{t=\underline{T}_{ij}}^{\overline{T}_{ij}} e_{jt}^2}} \tag{7}$$

Here e_{it} is the OLS estimate of the error term u_{it}, and the sample counterparts are $\hat{p}_{ij}$, which allows for pairwise cross-sectional correlation estimation. Where $\underline{T}_{ij}$ is the lower bound, and $\overline{T}_{ij}$ is the upper bound, which both indicate the overlap of time periods t over the cross-sections i and j. The number of time periods overlapping cross-sections is denoted by the test command CDTEST also provides the modified Breusch Pagan LM test results for large N and small T_{ij} and the Pesaran CD test.

In addition to the Breusch Pagan LM test, the data was tested for serial correlation with the ARIMA procedure of SAS (SAS Institute Inc., 2017). In this case, the residuals of the single cross-sections were tested for serial correlation. Since serial correlation could not be ruled out, the Newey-West was used (Newey & West, 1987). The Newey-West estimator is heteroskedasticity and autocorrelation consistent (HAC) and uses the Bartlett kernel and bandwidth determined by the equation based on the sample size and no adjustment to degrees of freedom and no prewhitening (SAS Institute, 2017).

Results

The data analysis is split into two parts. First, the monthly data is analyzed, and the quarterly next. Then, two models are tested for each time horizon with either the total quantity purchased or the contract length as the dependent variable.

Monthly Data

For the monthly data set, three additional traders had to be omitted from the analysis because of a lack of observations. As a result, the number of cross-sections, the combination of traders and delivery locations, for the monthly data set is 142. Time-sections in the data set are 76 individual months. Since not all traders have deliveries to each location each month, the data set is considered unbalanced. Averages for the most important variables can be found in table 1. The average number of bushels purchased via forward contract per trader and specific location per month is 413,735 bushels. An average contract length is 123 days from contract set-up to delivery. On average, the derived gateway basis is 0.014 USD per bushel, while the first lag of the average *TM* is about 0.1 USD per bushel.

For the first model, the total amount of bushels purchased via a forward contract by a trader for a specific location (*TBU*) serves as the dependent variable:

$$TBU_{ijt} = \beta_0 + \beta_1 AvgDerivedGWBasisMonth_{ijt} \\ +\beta_2 AvgTMMonth_{ijt-1} + \alpha_{ijt} + u_{ijt} \tag{8}$$

AvgDerivedGWBasisMonth is the average derived gateway basis that a trader could obtain the day of contract set-up, and $AvgTMMonth_{ijt-1}$ is the average *TM* a trader achieved for a specific location in the month before the purchase. The unobserved effect is a_{ijt} and u_{ijt} is the error component.

Several steps were taken to determine the most suitable panel data analysis method. While the Breusch Pagan LM test did not reveal any cross-sectional dependence, the ARIMA procedure could not out rule serial correlation. Since serial correlation cannot be ruled out, the Newey West heteroskedasticity- and autocorrelation robust estimator is applied for further calculations.

In the following step, a one-way fixed-effects analysis was conducted for fixed cross-sections. The F-test showed that the cross-sections are not showing fixed heterogeneity. The p-value of the one-way fixed-effects for fixed cross-sections is <.0001, which leads us to assume that the fixed effects model is the better choice than a pooled model. Next, the F-test was conducted again for the one-way fixed-effects model but this time for the time-sections. While the test result is insignificant with a p-value of 0.1958, the data structure is still more suitable for assuming fixed time effects. Wooldridge (2016) points out that fixed-effects can still be the better fit, mostly when each a_i can be thought of as a separate intercept, which applies in this case (P. 445).

To double-check if fixed effects were the right choice for the panel data, a two-way random-effects test was conducted to obtain a Hausmann-test m-value. Simultaneously, the test provided a p-value of 0.0110, which is larger than alpha at 0.5%, but it was decided to use a two-way fixed-effects model with the Newey-West estimator to better suit the data set.

Finally, a two-way fixed-effects model was used to test the null hypothesis of no influence on forward purchasing behavior by the current monthly average basis or the average TM achieved in the previous month per trader. The two-way fixed-effects model's F-test shows that the two-way fixed-effect is appropriate for the data with a p-value of 0.001, which is lower than alpha at 5%. Therefore, the null hypothesis was rejected in favor of the fixed-effects method.

The results of the panel data analysis for the monthly data for the total bushels purchased for a specific location are in table 2. For the monthly data, we find no evidence that the current average derived gateway basis significantly influences the quantity of bushels purchased by a trader on a forward contract. In addition, although of the expected sign, the average TM a trader achieved in the prior month has no significant impact on the quantity of bushels purchased by a trader in a basis contract. A negative coefficient, β_1, for $AvgDerivedGWBasisMonth_{ijt}$ implies a higher basis is associated with fewer bushels purchased on a forward contract. A high basis would indicate that the cash price of a commodity is relatively high in comparison to the respective futures price. While a high basis may be attractive to sellers it is not attractive for purchasers to cover their demand when cash prices are relatively high. Based on the relationship between the basis and the cash price it would not be surprising – in the absence of purchasing pressure – that commodity purchasers hesitate to buy corn when basis and cash prices are relatively high. Our finding that current basis does not impact forward contract size may be related to purchasing pressure. If the company should run low on stock and may have to temporarily halt feed mill production, and then even at relatively high basis levels, traders may be influenced by wanting to avoid such production halts. Since our data focuses on forward trades, last minute forced purchases are not included in our analysis.

A second measure of how comfortable a trader feels about a trade is the forward contract's length, meaning the sum of the days between contract set up and delivery. The model used for the total bushels forwarded is modified to test the derived gateway basis's influence and the previous average *TM* on the average contract length:

$$\begin{aligned} Contract\ Length_{ijt} &= \beta_0 + \beta_1 AvgDerivedGWBasisMonth_{ijt} \\ &+\beta_2 AvgTMMonth\ _{ijt-1} + \alpha_{ijt} + u_{ijt} \end{aligned} \tag{9}$$

For this second model with contract length as the dependent variable, the same data set used for the first model of total bushel purchased via forward contract was used again to test for dependent variable contract length. The model with the contract length dependent variable was chosen to be a two-way fixed-effect model with the Newey-West estimator after consulting the diagnostic tests.

The results for the contract length change for the monthly time horizon are in table 3. As explained previously the average derived gateway basis has a negative signal, as we would expect due to the higher cash price purchase level associated with higher basis levels. This time this coefficient is also significant. If the average derived gateway basis increases by one dollar per bushel, the contract length decreases by 107 days. A higher derived gateway basis, and higher commodity cash price levels means higher risk for the traders, leading the commercial traders to become more risk averse. The average *TM* the traders achieved in the month prior to setting up new contracts did not have a significant impact on the contract length.

Quarterly Data

For the quarterly data, the same models as for the monthly data were tested. As the same tests were consulted, the main variables are the same. Descriptive statistics for the contract length, the total quantity purchased via forward contracts, the average gateway basis, and the average *TM* in the quarter before a purchase can be found in table 4. The total quantity of bushels purchased via forward contract by a trader for a location increased in the quarterly period to 1,034,385 bushels. The average contract length for newly arranged forward contracts was about 90 days. The average derived gateway basis was 0.04 USD per bushel, and the average *TM* of a trader during the quarter before a purchase was 0.09 USD per bushel. Further, the quarterly data has 126 cross-sections of traders and delivery locations and 25 time-sections or quarters. The model for the panel data analysis for the quarterly data with the total quantity purchased via forward contract can be written as:

$$TBU_{ijt} = \beta_0 + \beta_1 AvgDerivedGWBasisQuarter_{ijt} + \beta_2 AvgTMQuarter_{ijt-1} + \alpha_{ijt} + u_{ijt} \quad (10)$$

Model 10 is the same as model 8, with the only difference in the time horizon being quarters rather than months. The results for model 10 are in table 5. Same as for the monthly data, several tests were conducted to find the best fitting model. An F-test result of the one-way fixed-effects model revealed a p-value of 0.0001 which leads to the decline of the assumption of the data's poolability and that cross-section effects are present. A similar result was obtained from the F-test of the one-way fixed time-effects model. Here a p-value of 0.0001 also indicated the presence of time-series effects. A final test to look at the Hausman test to decide between random- or fixed-effects indicates that random-effects would be favorable with a p-value of

0.288 for a two-way random-effects model. The better fit for the data is the Two-way fixed-effects approach given the data structure and the one-way fixed-effects F-test results.

As for the monthly test, a Breusch Pagan LM test (Breusch & Pagan, 1980) was consulted. The test's p-value leads to accepting the null hypothesis, assuming no proof of cross-sectional dependence in the data. The two-way fixed-effects panel analysis results were checked with the ARIMA procedure to check for autocorrelation on the time-series part of the data. Again, the trader's residuals were extracted from a two-way fixed-effects panel analysis and individually by trader tested with the ARIMA procedure. Since the tests did not out rule serial correlation, the Newey West estimator (Newey & West, 1987) is applied to the quarterly data analysis. The SAS FixTwo command is an LSDV model. The main drawback of the LSDV model is its restrictiveness with a large number of subjects. In the present case, the number of subjects should not cause any problems.

Like for the monthly data, neither the average derived gateway basis, nor the average *TM* are significant. To also test the contract length as a risk measure for a forward contract, model 9 was adjusted to accommodate the quarterly data and can now be written as:

$$Contract\ length_{ijt} = \beta_0 + \beta_1 AvgDerived\ GWBasisQuarter_{ijt} + \beta_2 AvgTMQuarter_{ijt-1} + \alpha_{ijt} + u_{ijt} \quad (11)$$

The results for model 11 provide evidence that the current derived gateway basis significantly impacts a forward contract's contract length and can be seen in table 6. For example, if the derived gateway basis increases by one dollar per bushel, the contract length declines by 98 days. This result is similar to the monthly data where the contract length was also shortened with an increased derived gateway basis.

Discussion

The data analysis results indicate that past performance has little or no influence on grain traders' decision-making process to use forward contracts to cover the demand for corn for the feed mills they supply. The loss aversion literature of Shefrin et al. (1985), Nofsinger (2017), and Coval & Shumway (2005) would have not expected to see an increase in bushels purchased via forward contracts when the past performance is good, here a high *TM*. It would be be difficult to attribute a possible increase in bushels purchased with forward contracts for trader with a high *TM* to the house money effect based on positive past performance. Our study is not set up in single, consecutive gambles as the horse race betting observed by (Suhonen & Saastamoinen, 2017). Their horse race experiment takes place over a short period. Neither is our study set up like the experiments by (Ackert, Charupat, Church, & Deaves, 2006; Suhonen & Saastamoinen, 2017), who set up experiments with students. Another significant difference is that the commercial grain dealers are not working with their private wealth.

On the other hand, the commercial grain traders may still feel optimistic about their performance and show more confident behavior when engaging in forward contracts. Our results are close to Friesen et al. (2021), who found that past gains from repurchase decisions had almost no evidence to influence future repurchasing decisions. The past performance of the grain traders did not have significant impact on the contract length of future forward purchases. We would have expected for previously successful traders to engage shorter, less risky, forward purchases on the short time horizon.

The derived gateway basis at the time of purchase had a negative influence on the length of contracts. Traders would prefer a low derived gateway basis, which means that the cost of corn at the time of purchase in the physical market would be relatively low compared to the futures market. If the derived gateway basis is high, traders might also cover the demand for corn in the local spot market at the time of need rather than purchasing in advance. The derived gateway basis has a significant influence on the contract length. If the derived gateway basis is high, traders shy away from engaging with contracts with delivery times far out in the future. Limiting purchases when the derived gateway basis is comparatively high gives the traders the chance to purchase corn in the spot market or with shorter forward contracts at a time when the derived gateway basis is lower. While not covering their demand with forward contracts, the traders expose themselves to the risk that the derived gateway basis of the spot price when the corn is needed may be even higher.

The risk components of forward contracts are the volume and contract length. If either the size or the length of the contract increases, so does the attached risk. It can be assumed that if a trader agrees on a contract for delivery in the future, the larger the volume and the longer the contract length, the more comfortable they are that the contract agreement is a good deal for them.

Recommendations

In our case, the commercial grain traders of this firm did not show any significant influence in their froward trading behavior based on their past performance. We are unaware of an internal policy of a firm-specific reward system that would reward traders for their trading performance in terms of the margins achieved in previous periods. It appears that traders are more influenced by the derived gateway basis and, therefore, the price level at contract set up, which would not support a bonus system based on past performance. While the firm indeed is interested in the profitability of its operations, there might be higher importance of the risk management aspect of the traders' behavior. Meeting the demand of the individual feed mills is of utmost importance to the firm to avoid production stops due to supply chain issues based on poor purchasing practices of input commodities. Feed optimization can increase the profitability of meat-producing companies, which may lead to a high interest in reducing the risk in feed production rather than achieving the highest possible profitability when purchasing inputs. A stable meat and poultry industry are also in the interest of the American consumer. The U.S. meat and poultry industry provided 65.2 pounds of chicken and 54.6 pounds of beef per capita in 2018 (U.S. Department of Agriculture, Economic Research Service, 2018) while also employing about 5.4 million people while the poultry industry alone provided 2.1 million jobs in 2020 (Venable, G., 2021).

The results highlight the importance of basis trading. We suggest that traders who understand the local spot markets and their relations with futures markets use this information heavily in their grain purchasing decision-making process. We recommend that the education of commercial grain traders and other market participants should focus on the understanding of basis trading and risk mitigation strategies rather than only focusing on profitability. The trader's past performance does not heavily influence the traders from our dataset in their decision making, showing the importance of risk management in the agricultural- and food industry commodity procurement over only focusing on high margins. Basis trading is not only crucial for the traders purchasing grain but also for the farmers to make better selling decisions.

Limitations

The number of traders is somewhat limited in the data set used for this study. A larger sample of commercial grain traders, ideally from several different companies, would be beneficial for the analysis. Additionally, the traders in this data set are required to cover the demand their company dictates. Any demand of input commodities not covered with forward contract purchases forces the traders to buy in the spot market at the spot price, not interrupting their operations' production process. Meeting the demand and considering the feed mills' storage capacities may impact the traders' decision-making and profitability, especially when comparing different locations. Also related to the limited number of traders are the time spans selected for this study. It could be possible that commercial traders would be influenced by their previous performance if the time periods were shorter, weekly or ideally even daily. Our fairly long monthly and quarterly periods may lead to traders not remembering their past performance. Coval & Shumway (2005) were able to observe loss aversion based on trader performance by dividing daily performance in mornings and afternoons.

A primary distinction between the traders of our dataset and the main body of behavioral economics financial literature is that the traders are not using their private wealth for commercial transactions. Individuals using their personal wealth in transactions may be more sensitive to changes in wealth than employees in a work setting.

Conclusion

Looking at two possible reference points for commercial grain traders to base their trading decisions on, it seems that the current derived gateway basis is more important in influencing trading decisions than the previous performance of the individual traders. But even the derived gateway basis only influences the risk a trader is willing to engage in in terms of the duration of the forward contracts and does not influence the amount of bushels the traders purchase via forward contracts. Since our dataset is limited, it would be beneficial to test our models with more traders and possibly shorter time periods than months to see if a more recent previous performance impacts the decision-making of commercia grain traders.

References

Ackert, L., Charupat, N., Church, B., & Deaves, R. (2006). An experimental examination of the house money effect in a multi-period setting. *Experimental Economics, 9*(1), 5-16. doi:10.1007/s10683-006-1467-1

Ajmani, V. B. (2009). *Applied econometrics using the SAS system*. Hoboken, N.J: John Wiley. Retrieved from http://www.loc.gov/catdir/enhancements/fy0810/2008004315-d.html

Baltagi, B. H. (2008). *Econometric analysis of panel data*. Chichester, UK: John Whiley & Sons Ltd.

Beckert, J. (2009). The social order of markets. Theory and Society, 38(3), 245-269. doi:10.1007/s11186-008-9082-0

Breusch, T. S., & Pagan, A. R. (1980). The lagrange multiplier test and its applications to model specification in econometrics. *The Review of Economic Studies, 47*(1), 239-253. doi:10.2307/2297111

Coakley, J., Dollery, J., & Kellard, N. (2011). Long memory and structural breaks in commodity futures markets. Journal of Futures Markets, 31(11), 1076-1113.

Coval, J. D., & Shumway, T. (2005). Do behavioral biases affect prices? The Journal of Finance, 60(1), 1-34.

Fama, E. F. (1970). Efficient capital markets: A review of theory and and empirical work. The Journal of Finance, 25(2)

Friesen, G. C., Jeutand, N. P., & Unlu, E. (2021). The effect of unsuccessful past repurchases on future repurchasing decisions. *Management Science,*

Greene, W. H. (2003). *Econometric analysis* (Eighth edition, global edition ed.). New Jersey: Prentice Hall. Retrieved from http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&local_base=BVB01&doc_number=029693716&sequence=000002&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica, 46*(6), 1251-1271. doi:10.2307/1913827

Holst, G. S., Hermann, D., & Musshoff, O. (2015). Anchoring effects in an experimental auction–Are farmers anchored?. Journal of Economic Psychology, 48, 106-117.

Hsiao, C. (2007). Panel data analysis—advantages and challenges. *Test, 16*(1), 1-22. doi:10.1007/s11749-007-0046-x

Jaba, E., Robu, I., & Balan, C. B. (2017). Panel data analysis applied in financial performance assessment. Revista Română De Statistică, 65(2), 3-20.

Kahneman, D. (2011). *Thinking fast and slow*. New York: Farrar, Straus and Giroux.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision making under risk. *Econometrica, 47*(2)

Kahneman, D., & Tversky, A. (1984). Choices, values and frames. *American Psychologist, 39*, 341-350.

Lorton, S. & White, D. (2002) Merchant's Edge. Champaign, IL: Stipes Publishing.

Martin, H. M., Smith, K. T., & Murphy Smith, L. (2016) A financial and public policy analysis of the meat industry. International Journal of Economics and Accounting, 7(1).74-88.

Mattos, F. L., & Fryza, S. A. (2014). Do farmers exhibit disposition effect? evidence from grain markets. *Managerial Finance, 40*(5), 487-505. doi:10.1108/MF-07-2012-0157

Mattos, F. L., & Zinn, J. (2016). Formation and adaptation of reference prices in grain marketing: An experimental study. *Agricultural Economics, 47*(6), 621-632. doi:10.1111/agec.12260

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica, 55*(3), 703-708. doi:10.2307/1913610

Nofsinger, J. R. (2018). Behavioral aspects of commodity markets. In K. H. Baker, G. Filbeck & J. H. Harris (Eds.), Oxford University Press.

Odean, T. (1998). Are investors reluctant to realize their losses? *The Journal of Finance, 53*(5), 1775-1798. doi:10.1111/0022-1082.00072

SAS Institute. (2017). *SAS/ETS 14.3 user's guide the PANEL procedure.* (). Cary, NC: SAS Inistiute INC.

SAS Institute Inc. (2017). *SAS/ETS 14.3 user's guide the ARIMA procedure*. Cary, NY, USA: SAS Inistiute INC.

Shefrin, H. (2006). *Behavioral corporate finance*. New York: McGraw-Hill.

Shefrin, H. M., & Statman, M. (1985). The disposition to sell winners too early and ride losers too long. *The Journal of Finance, 40*(3), 777-790. Retrieved from http://www.econis.eu/PPNSET?PPN=258853220

Suhonen, N., & Saastamoinen, J. (2017). How do prior gains and losses affect subsequent risk taking? new evidence from individual-level horse race bets. *Management Science, 64*(6), 2797-2808. doi:10.1287/mnsc.2016.2679

Shleifer, A. (2000). Inefficient markets (1. publ. ed.). Oxford [u.a.]: Oxford Univ. Press.

Thaler, R. (1980). Toward a positive theory of consumer choice. Journal of Economic Behavior and Organization, 1(1), 39-60. doi:10.1016/0167-2681(80)90051-7

Thaler, R. H., & Johnson, E. J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. Management Science, 36(6), 643-660.

Thaler, R. H., Tversky, A., Kahneman, D., & Schwartz, A. (1997). The effect of myopia and loss aversion on risk taking: An experimental test. The Quarterly Journal of Economics, 112(2), 647-661

U.S. Department of Agriculture. Economic Research Service. (2018). Food Availability and Consumption. Retrieved from: https://www.ers.usda.gov/data-products/ag-and-food-statistics-charting-the-essentials/food-availability-and-consumption/

Venable, G. (2021). The economic impact of the poultry industry. Poultry Perspectives. The National Provisioner. Retrieved from: https://www.provisioneronline.com/articles/110609-the-economic-impact-of-the-poultry-industry

Wooldridge, J. M. (2016). Introductory econometrics (Sixth edition, student edition ed.). Boston, MA: Cengage Learning.

.

Tables

Table 1

Descriptive statistics of the main variables in the monthly dataset

Variable	N	Mean	Std Dev	Minimum	Maximum
TBU (Total bushel purchased)	2,298	413,735	379,070	942	3,201,306
Contract Length	2,298	123.02	95.94	7.00	608.50
Average Derived Gateway Basis	2,298	0.01	0.37	-0.87	1.60
Average Margin t-1	2,298	0.10	0.38	-5.33	10.12

Notes: The contract length is in days. Derived gateway basis means the basis (price difference between spot and futures market) was adjusted for fright or backed off to a gateway. Average margin lag 1 represents the average margin a trader achieved in the most recent period.

Table 2

Monthly Fix-Two panel data model for total bushels purchased per trader for a specific location from August 2012 to February 2018

Intercept	8,802
	(105,172)
Avg DerivedGWBasis Month	-32,663
	(27,053)
Avg Margin Month t-1	14,577
	(13,127)
R2	0.58
Num. of Cross Sections	142
Times Series Length	76

Notes: ***p<0.01, **p<0.05, *p<0.10. Standard Errors are in parentheses. The model was estimated with two-way fixed-effects panel data analysis.

Table 3

Monthly Fix-Two panel data model for contract length per trader for a specific location from August 2012 to February 2018

Intercept	190.03***
	(29.64)
Avg DerivedGWBasis Month	-106.64***
	(8.75)
Avg Margin Month t-1	-1.93
	(5.53)
R2	0.48
Num. of Cross Sections	142
Times Series Length	76

Notes: ***p<0.01, **p<0.05, *p<0.10. Standard Errors are in parentheses. The model was estimated with two-way fixed-effects panel data analysis.

Table 4

Descriptive statistics of the main variables in the quarterly dataset

Variable	N	Mean	Std Dev	Minimum	Maximum
TBU (Total bushel purchased)	813	1,034,385	1,321,626	1,000	9,046,667
Contract Length	813	89.60	73.80	7.00	372.00
Average Derived Gateway Basis	813	0.04	0.36	-0.89	1.52
Average Margin t-1	813	0.09	0.22	-0.55	2.02

Notes: The contract length is in days. Derived gateway basis means the basis (price difference between spot and futures market) was adjusted for fright or backed off to a gateway. Average margin lag 1 represents the average margin a trader achieved in the most recent period.

Table 5

Quarterly Fix-Two panel data model for total bushels purchased per trader for a specific location from August 2012 to February 2018

Intercept	-400,529
	(563,159)
Avg DerivedGWBasis Quarter	-200,958
	(171,713)
Avg Margin Quarter t-1	46,636
	(214,204)
R2	0.45
Num. of Cross Sections	126
Times Series Length	25

Notes: ***p<0.01, **p<0.05, *p<0.10. Standard Errors are in parentheses. The model was estimated with two-way fixed-effects panel data analysis.

Table 6

Quarterly Fix-Two panel data model for contract length per trader for a specific location from from August 2012 to February 2018

Intercept	4.64
	(30.39)
Avg DerivedGWBasis Quarter	-98.05***
	(10.17)
Avg Margin Quarter t-1	7.91
	(20.65)
R2	0.49
Num. of Cross Sections	126
Times Series Length	25

Notes: ***p<0.01, **p<0.05, *p<0.10. Standard Errors are in parentheses. The model was estimated with two-way fixed-effects panel data analysis.

The Influence of the World Agricultural Supply and Demand Estimate Report Estimates on Corn Purchasing and Marketing Behavior

Abstract

The World Agricultural Supply and Demand Estimates (WASDE) published by the USDA are of great importance for agricultural markets. Due to the great importance of information in agricultural markets private forecasts are also available at a cost prior to the WASDE release. This study further investigates the importance of the monthly publication for commercial corn traders and corn producers in their purchasing and selling decisions throughout the corn marketing year compared to the private forecasts. The data is organized in weekly periods and grouped by seasons. First, we take a look at the importance of the difference in settle price on the purchasing and selling decisions of corn traders. Next, the role of the marketing week is investigated. A seven-day period prior to the WASDE release, a seven-day period starting with the WASDE release, and a seven-day period following the WASDE release week are created to gain insights into trading activity around the WASDE release. Finally, the changes form the week pre WASDE release to the WASDE week are studied to lean more about the importance of the WASDE report compared to private forecasts. Our results reveal differences in the use of information by commercial corn traders compared to corn producers when making marketing decisions.

Keywords: commodity markets; Dept. of Agriculture; commercial grain traders; agricultural producers

Introduction

Futures and spot markets for storable commodities depend heavily on current market information for price discovery purposes. Accordingly, the United States Agriculture Department (USDA) has been the primary agricultural markets data collector. For more than 150 years, trusted public data and reports have been published by USDA (Isengildina-Massa, Karali, & Irwin, 2020). Two thousand one hundred publications are currently provided by the USDA Economics, Statistics, and Market Information System (ESMIS).

One of the most anticipated periodical reports is the World Supply and Demand Estimates (WASDE), prepared by the World Agricultural Outlook Board (WAOB). This monthly report provides information and forecasts for the annual production, trade, and use of the major agricultural commodities in the U.S. and worldwide. Isengildina-Massa, Karali, & Irwin (2013) point out the public good role of the WADSE. The report's high quality and "benchmark" role can be explained by the problem of private underinvestment in information and the invested resources of the USDA (p.5086). The WASDE report is credited to "enhance the efficiency of resource allocation in the agricultural sector (Adjemian, 2012, P. 238). The importance of the report is highlighted by the fact that its release influences market prices. For example, Isengildina-Massa, Irwin, Good, & Gomez (2008) found that the return variance on report release days is 7.4 times larger for corn than on other days. Due to the importance of the information provided, the WASDE is prepared in a secure environment, preventing information from leaking. Therefore, on the scheduled release dates, interested parties can access the report at the exact scheduled release date (Adjemian, 2012).

Nevertheless, the role of the WASDE has been questioned since the 1980s based on a mixture of cuts in fiscal spending and the rise of private information sources (Just, 1983; Xie, Isengildina-Massa, Dwyer, & Sharp, 2016). Besides publicly-funded reports, private forecasts are typically available to paying subscribers. The popular press is usually quick to share information from those private forecasts with a broader audience. Forecast reports by private companies preceed the USDA in their report publications by a couple of days, raising the question of whether the private information is making the costly public WASDE announcements somewhat redundant.

In 2013 the October WASDE report was missing due to a lapse in appropriations (Adjemianet al., 2018). The importance of the report as a market signal was highlighted by Adjemian et al. (2018), who found that in the absence of the missing WASDE report of 2013, futures and options prices did not react in a typical manner the day the report should have been released. In addition, Egelkraut et al. (2003) found that private forecasts did not provide superior crop production forecasts compared to the WASDE. As such, the private forecasts do not supplant USDA forecasts. They concluded that while the economic value of the private and public reports appears to be similar, private reports could provide a small window of opportunity before the public absorbs the information. Our study has the unique opportunity to study the effect of private information and public report releases on commercial trader behavior in a private company. The data also contains sales initiated by corn producers trading with the private company, allowing a view on the producer's behavior.

Literature

The robust version of the efficient market theory states that prices should reflect all public and private current information at any time, implying that rational agents update their price expectations immediately upon receiving any new information. Numerous studies have investigated commodity market responses to the different USDA reports. The USDA's most significant monthly report is the WASDE, which projects national and international supply and demand factors across all agricultural markets. Besides the WASDE, the USDA issues a national report known as the Crop Progress[2] report and the Crop Production Report[3]. The Crop Progress report observes crops' development status through the growing season, and the Crop Production report projects production and yields.

Lehecka (2014) found that the USDA Crop Progress report "provides additional and valuable information for participants in corn and soybean markets," despite private sector analysis and available weather information (p.103). McKenzie (2008) summarized from previous studies that crop reports produced by private companies "are good proxies for market expectations" and that private reports anticipated "at least some information contained in USDA reports" (p.352). Previous research has focused on a broad aggregate market trader/agent behavior and has assessed futures price responses to the information contained in private and USDA reports. The extant literature shows that futures prices respond to new or unanticipated information contained in USDA reports. This unanticipated or surprise element of the report is measured as the difference between official USDA numbers of grain stocks and private anlalyst forecasts of grain stocks. When USDA grain stock figures are larger than private analyst forecasts – indicating supply is larger than previously expected by the market – prices fall, and vice versa. Unlike prior studies using aggregate (future market) data, our research will differ by

---

2 More information about the Crop Progress can be found here: http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1048

3 More information about the Crop Report can be found here: http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1046 analyzing the impact of private forecasts and USDA reports to influence producer selling behavior or private company purchasing behavior, unlike prior studies using aggregate (future market) data.

Behavioral economic heuristics may play a role in marketing and purchasing of commodities around the release of USDA reports. The foundation for extensive research in anchoring, or the use of reference points, was laid out by Tversky and Kahneman (1974) when they introduced the heuristic principles people use when faced with decisions that require the assessment of probability. “When faced with uncertainty, people will grasp at straws to find a basis for the view” (Montier, 2002, p.5). Unfortunately, the use of those heuristics tends to create systematic errors. Anchoring is described by Tversky and Kahneman (1974) as different estimates yielded by different starting points leaning towards the initial values. Later Kahneman describes anchor effects as: “Cases in which a stimulus or a message that is designated as irrelevant and uninformative nevertheless increases the normality of a possible outcome” (Kahneman, 1992, p.308). Anchors may be more or less strong depending on the functioning of norms and an individuals’ impression of relevant values and the stimuli’ level of judgment. Anchoring is the process of adjusting estimations away from an initial starting point, which serves as an anchor, where adjustments are biased towards the initial starting value. Hence, different anchors lead to different estimations (Tversky & Kahneman, 1974).

In contrast, the reference point plays an essential role in determining the evaluation of gains and losses, which in Bernoulli’s theory would have been the state of wealth to determine utility. Typically reference points are used as “salient neutral points on evaluation scales” for the judgment of abrupt changes “at which the slope of the value function shows a sharp transition” (Kahneman, 1992, p. 310). Both reference points and anchors can be adjusted over time.

The literature on anchoring and reference price effects in agricultural commodities marketing decisions is somewhat limited. Mattos and Zinn conducted one of the very few studies with farmers about the formation and adaption of reference prices in 2016 in Canada. Seventy-five farmers participated in a pen and paper experiment about marketing their grain over ten months. The researchers found that farmers used reference prices (which they called satisfy price to express their satisfaction when achieving this price) and regularly updated those reference prices. Contrary to falling market prices, farmers set their reference prices above the market price and then updated faster when the market prices rose. The reference price was also updated based on the farmer's future market price expectations, when price trends changed, or when the farmer's price expectation had changed (Mattos & Zinn, 2016). The experiment also showed that farmers tend to sell proportionally more grain when the current market price is above their reference price and vice versa. Problematic is the finding that farmers sold less grain when they expected a bearish market, which might lead them to hold on to their grain for too long to achieve a satisfying price for the grain.

Studies about anchoring or reference prices are more prevalent in other economic research disciplines than agriculture. Several studies in the financial literature found an inconsistency in the efficient market theory and its underlying assumptions. In their 2004 study, *the 52-Week High and Momentum Investing*, George and Hwang (George & Hwang, 2004) found that the 52-week high price stock price strongly influenced how traders react to new information regarding the particular stock. If the price was near the 52-week high, traders hesitated to bid on a higher price, even though good news indicated a higher value. Once the excellent news accumulated, traders moved the price up continuously. In the opposite case, when the price was far below the 52-week high, traders tended to hold on to the stock. They describe this initial hesitation to adjust to new information initially as an under-reaction and as an anchoring bias towards the 52-week high price that the market resolves in the long run if information prevails. The authors concluded that the 52-week high price as an anchor was a better predictor for future returns than the historical returns would have been (George & Hwang, 2004).

The 52-week high price might as well play an essential role as an anchor for farmers. Farmers can monitor the commodity prices in real-time online, subscribe to newsletters or specialists, or communicate with their marketing business partners and base. Zinn and Mattos (2016) found that the highest price to date was statistically significant in their experiment in the creation of reference, also called satisfy prices in the experiment. Grinblatt & Keloharju (2001) found that monthly high or monthly low stock prices influenced stocks' buy or sell decisions. On the contrary, Baucells, Weber, and Welfens (2011) found the influence of historical peaks only marginal. They described reference price formation as depending on the purchase price of the stock and the latest available price. Kahneman (1992) highlighted the importance of future research to gain deeper insights into the formation and combination of different reference prices.

Reference points are used to compare outcomes to a reference, which can be for example a comparison of a price to an expected price. In decision sciences, reference points are often considered to be created in conjunction with the expectations of decision-makers. Further reference points are often subject to change and updates when the underlying expectation is adjusted to new information (Tonsor, 2018; Baucells et al., 2011; Wenner (2015). Anchoring bias describes how decision-makers are influenced by salient information or first impression in the decision maker's environment that are used as anchors to make decisions. A typical example here is to ask participants to make guesses after spinning a wheel that will show randomly small or large numbers. Participants that are shown higher numbers on the wheel tend to make higher guesses for unrelated questions than participants being shown small numbers on the random wheel (Epley, 2013, Anchoring, p.28). Extensive research has been done on reference points and consumer decision-making. However, much less research has concerned reference point decision-making in agricultural commodity markets. For example, how producers in agriculture use reference points in decision-making under uncertainty has received little attention (Tonsor, 2018), and no known research has analyzed this issue beyond the farmgate. A study from 2016 by Bianchi, Drew, & Fan found assertive anchoring behavior when using the 52-week high momentum strategy in commodities to learn about conservatism bias. Conservatism bias was first described in the economic field by Barberis, Shleifer, & Vishny in 1998 and is the delay in updating beliefs and slow reaction to new evidence. Besides the strong anchoring, Bianchi et al. (1916) confirmed conservatism bias for commodity trading. In our case, since the commodity traders and corn producers are not traditional financial investors, we use the stock-to-use ratio predictions of the monthly USDA WASDE report as a reference. The WASDE reports are highly anticipated populations in the physical commodity trading world as the stocks-to-use ratio and other estimates in the report are used to form opinions about supply and demand situations for the traders and the producers. It is a common belief that momentum trading or price rally trading exists when farmers react to rising commodity prices and sell greater quantities when commodity prices rise. First, the futures prices rise, then the spot prices follow. This momentum continues until spot prices catch up and the basis decreases. In our study, we test both the release of the WASDE report with trading activity in the weeks during, at release, and after release and the surprise of the stocks-to-use ratio publication of the WASDE after private forecasts as reference points for corn producers and commercial traders.

Data

A private grain trading company provided the data set for this study. All corn trading transactions between July 2011 and February 2018 are recorded in the data. The data was set up in weekly periods and divided into four seasons. In the next step, the data set was split into two parts, one for the traders and the other for the producers, which are later analyzed separately. The first group are the transactions entered into by commercial traders. The commercial traders initiate their buying decision using contracts to purchase corn with grain merchandising firms.

In contrast, with regard to our second group consisting of producers, the producers initiate the timing of the marketing/selling decisions. Therefore, the split data set will allow two different analyses based on a sale or purchase decision-maker. This separation is important as the decisions made by corn producers differ from the commercial grain traders, and we want to identify the factors that impact these buying and selling decisions or behavior. Units in the dataset are in U.S. dollars (USD) per bushel.

In addition, WASDE surprise information and nearby (contracts closest to maturity time) corn futures settle prices were collected over the same period. The settle price refers to the last trading day price for the nearby corn futures on the futures market exchange. The *WASDE Surprise* is calculated by comparing the stocks-to-use ratio of the WASDE outlook to the stocks-to-use ratio forecasted by private firms. The private forecasts were obtained from Bloomberg. First, the beginning stock and total production need to be estimated, then the total use of corn can be subtracted from the available corn to calculate the ending stock per year. The stocks-to-use ratio is the percentage of corn that remains in storage as a carryover divided by the total use of corn in the United States for a corn marketing year. The *WASDE Surprise* is the percentage difference between the WASDE numbers and the private forecast numbers.

The WASDE and private forecast estimates for stocks and use are updated monthly. The carryover stock estimates influence the supply and demand market situation that determines the price for commodities. The median *WASDE Surprise* is the natural logarithm of the WASDE stocks-to-use ratio minus the natural logarithm of the private forecast stocks-to-use ratio. Positive surprises reflect an unexpectedly larger supply of corn than the market anticipated and typically result in futures price decreases. In contrast, negative surprises reflect an unanticipated smaller supply of corn and are associated with futures price increases. The larger the surprise (positive or negative), the greater the likelihood that there was a lot of market uncertainty as to the true supply and demand picture prior to a release date. Hence, we use the size of surprises to proxy market uncertainty and hypothesize that large surprises and uncertainty may negatively impact commercial traders' purchasing behavior and producers' marketing behavior.

Several calculations had to be conducted to finalize the data preparation. All calculations are across all traders and all locations, or respectively across all farmers and all locations. First, for each weekly period, the total quantity purchased was calculated. All purchased quantities are the quantities the purchaser and seller agreed upon when the contract was initiated. Ultimately there is a possibility that not all the amounts contracted were delivered at contract maturity. Since our focus is on decision-making, we are looking at the agreed-upon quantity at the time of contract agreement. Secondly, the mean settle price of the nearby futures contract for corn per period was calculated and added. Both the total purchased quantity for the traders and the total sold quantity for the farmers and the mean settle price were log-transformed, and the difference of each to the previous period was recorded. They were respectively named *Diff Qty* and *Diff Settle*.

Dummy variables are created to categorize the dates on which transactions took place in weekly data. The category "*WASDE*" "starts with a WASDE release day and contains an additional six days. Category "*PreWASDE*" "contains the seven days before a WASDE release day, and finally, the last seven-day period "*PostWASDE*" "follows the WASDE release period. For example, in January 2013, the WASDE report was released on January 11. The *PreWASDE* period ranged from the 4th to the 10th, the *WASDE* period from the 11th to the 17th, and the *PostWASDE* period from the 18th to the 25th. Additional to the dummy variables for the different trade periods regarding the WASDE release, seasonal dummies were created. The WASDE report was released at 8.30 am EDT before January 1, 2013, and at 12 pm EDT after January 1 (USDA, 2012). This change in release time does not affect our data. Both times are before the futures trading for corn futures closes, and the daily settle price is registered at 1.20 pm C.T. (CME Group).

Seasonality is important for the marketing cycle of agricultural commodities, and to capture this impact dummy variables were created for *Fall* (September 1 to November 30), *Winter* (December 1 to February 28), *Spring* (March 1 to May 31), and *Summer* (June 1 to August 31). Each of the seasons includes three months per year.

In addition to the dummy variables, several interaction terms were created to capture specific effects for the seasons and weeks. First, the difference in settle price was paired with the seasons, (*Diff Settle* x *winter*, *Diff Settle* x *spring*, *Diff Settle* x *Summer*), and for the week types, (*Diff Settle* x *WASDE*, *Diff Settle* x *PostWASDE*). The final interaction term was the *WASDE Surprise* and the seasons (*WASDE Surprise* x *Winter*, *WASDE Surprise* x *Spring*, *WASDE Surprise* x *Summer*).

An interesting occurrence happened in October of 2013 when a lapse in federal funding hindered the release of a month's WASDE report (USDA announces cancellation and postponement of selected reports impacted by the lapse in federal funding.). As a result, the missing month was removed from the data.

Methods

Several separate analyses were conducted for this study. The data analysis relied on spreadsheets constructed using Microsoft Excel (Microsoft). In addition, all statistical tests were conducted with (SAS Institute Inc., 2019)[4].

The first model analyzes the change in the quantity of corn purchased by commercial grain traders employed by a large agribusiness firm based on the change in the average weekly futures settle price of corn. Second, the model was used to determine if weekly futures price changes affect the quantity of corn marketed/sold by farmers to the large agribusiness firm. *A priori*, we would expect positive price changes to incentivize producers to market their corn. Therefore, we hypothesize that weekly futures price increases will lead to increases in the quantity of corn sold by producers. At first blush, one might expect weekly futures price increases to negatively impact the quantity of corn purchased by commercial traders. However, given that our commercial traders are incentivized to buy corn on basis rather than price, and in addition that grain merchandising firms/elevators often lower their basis quotes in rising futures markets, we hypothesize that weekly futures price increases may also lead to higher quantities of

4 The data analysis for this paper was generated using SAS software. Copyright © [2002-2012] SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA.

corn purchased by our commercial trading group. The model was applied separately to the commercial grain trader group and the selling farmer group. Model 1 can be stated formally as:

$$Diff\ Qty_{it} = \beta_0 + \beta_1 Diff\ Settle_t + \beta_2 D_{2t} + \beta_3(Diff\ Settle_t\ x\ D_{2t})$$
$$\beta_4 D_{3t} + \beta_5(Diff\ Settle_t\ x\ D_{3t}) + \beta_6 D_{4t} + \beta_7(Diff\ Settle_t\ x\ D_{4t}) + \varepsilon_{it} \qquad 1$$

Where $_i$ stands for either the commercial grain traders or the selling farmers, and $_t$ represents the time period. The weekly quantity changes in corn either bought (by traders) or sold (by producers) serves as the dependent variable. The model's independent variables are the weekly average difference in corn futures settle price and the seasons. The different seasons are represented as indicator variables. The fall season serves as the base case season and is captured in the intercept term β_0. The following terms are used to represent the seasonal indicator or dummy variables:

D_2 = 1 if *Winter*, 0 if otherwise,
D_3 = 1 if *Spring*, 0 if otherwise,
D_4 = 1 if *Summer*, 0 if otherwise.

Further, we use three interaction terms to capture the potential impact of price changes during specific seasons: ($Diff\ Settle_t\ x\ D_{2t}$), ($Diff\ Settle_t\ x\ D_{3t}$), $Diff\ Settle_t\ x\ D_{4t}$). The error term is ε_{it}. Descriptive statistics for model 1 are in table 1 for traders and producers.

In the next model, model 2, the different weekly time periods around the release of the WASDE report were included to isolate the impact of the WASDE release as a factor affecting purchasing or selling behavior. Model 2;

$$Diff\ Qty_{it} = \beta_0 + \beta_1(Diff\ Settle_t) + \beta_2 D_{5t} + \beta_3(Diff\ Settle_t\ x\ D_{5t})$$
$$+\beta_4 D_{6t} + \beta_5(Diff\ Settle_t\ x\ D_{6t}) + \varepsilon_{it} \qquad 2$$

in model 2 D_5 and D_6 represent the different weekly periods around the WASDE release dates:

D_5 = 1 if *WASDE* week, 0 if otherwise,

D_6 = 1 if *PostWASDE* week, 0 if otherwise.

As in model 1, interaction terms were used ($Diff\ Settle_t\ x\ D_{5t}$), ($Diff\ Settle_t\ x\ D_{6t}$) to capture price impacts during specific time periods. In this case, the interaction terms capture price impacts during the week immediately following the report release (*WASDE)*, and price impacts during the week occurring two weeks following the release date (*PostWASDE)*. The *PreWASDE* period serves as the base category and is captured in the intercept term β_0 in equation (2), Descriptive statistics for model 2 can be found in table 2 for the traders and table 3 for the corn producers.

A third model was tested to see if the seasons, or the period when the report was released, affected the results. Model 3 was tested only for a dataset containing prices and quantities of corn purchased and sold immediately around the *WASDE*, starting with report release dates. Specifically, model 3 looks at the quantity changes from the *PreWASDE* week to the *WASDE* week. The dependent variable of model 3 measures the change in the quantity of corn sold or purchased from the week prior to the *WASDE* release to the immediate week following the *WASDE* release. In this sense, we isolate the specific impact of the report on trader purchasing and producer marketing behavior. The dependent variable is named "*WASDE Diff Qty*."

$$\begin{gathered} WASDE\ Diff\ Qty_{it} = \beta_0 + \beta_1\ WASDE\ Diff\ Settle_t + \beta_2 WASDE\ Surprise_t \\ + \beta_3 D_{2t} + \beta_4 (WASDE\ Diff\ Settle_t\ x\ D_{2t}) + \beta_5\ (WASDE\ Surprise_t\ x\ D_{2t}) \\ + \beta_6 D_{3t} + \beta_7 (WASDE\ Diff\ Settle_t\ x\ D_{3t}) + \beta_8 (WASDE\ Surprise_t\ x\ D_{3t}) \\ + \beta_9 D_{4t} + \beta_{10} (WASDE\ Diff\ Settle_t\ x\ D_{4t}) + \beta_{11} (WASDE\ Surprise_t\ x\ D_{4t}) + \varepsilon_{it} \end{gathered} \quad 3$$

As previously, indicator variables are used to capture seasonal effects, with the fall period representing the base period, which is captured in the intercept term β_0. In equation (3).

D_2= 1 if W*inter*, 0 if otherwise,

D_3 = 1 if *Spring*, 0 if otherwise,

D_4 = 1 if *Summer*, 0 if otherwise.

Further, we use three interaction terms of the continuous *WASDE Diff Settle* and each of the seasons ($WASDE\ Diff\ Settle_t$ x D_{2t}), ($WASDE\ Diff\ Settle_t$ x D_{3t}), ($WASDE\ Diff\ Settle_t$ x D_{4t}), For model 3, interaction terms were also used for the *WASDE Surprise* ($WASDE\ Surprise_t$ x D_{2t}), ($WASDE\ Surprise_t$ x D_{3t}), ($WASDE\ Surprise_t$ x D_{4t}). Table 4 represents the descriptive statistics for the commercial grain traders, and table 5 shows the descriptive statistics for the corn producers.

The SAS AUTOREG procedure was used to estimate all three model specifications. The AUTOREG procedure provides a battery of residual diagnostic tests to account for non-normality and serial correlation. Residual diagnostic results are presented at the foot of each results table and show that our model specifications do not suffer from non-normality, serial correlation, or heteroskedasticity. Next, further residual diagnostic checks were conducted with the SAS ARIMA procedure, primarily focusing on the autocorrelation check for white noise. Finally, the SAS MODEL procedure was consulted to rule out heteroskedasticity with the Breusch-Pagan test and White's test. Based on the residual diagnostics, the most suitable model was chosen for traders and producers and each of the three tested models. The Ordinary Least Squares (OLS) regression analysis requires that several assumptions are met. With the independence of the errors being the primary assumption. Since the problem of error dependence is not uncommon in time series analysis, the SAS AUTOREG procedure offers methods for autoregressive error correction. When a second-order autoregressive progress is selected in SAS, the Yule-Walker method of estimation, as described in Gallant and Goebel (1976), is applied rather than the OLS method.

Results

The analysis section is split into two separate parts to capture the different drivers for the decision-making of commercial grain traders and corn producers. First, the data for the commercial grain traders are analyzed, and the second section will investigate the farmer's behavior.

Commercial Grain Traders

This section evaluates the data for the commercial grain traders. The null hypothesis states that the release of the WASDE does not affect the purchasing behavior of commercial grain traders purchasing corn.

Model 1 commercial grain traders

First, model 1 is tested to assess if the weekly change in the average corn futures settle price influences the quantity of corn purchased by traders across seasons. The parameter estimates for model 1 for the commercial grain traders can be found in table 6. Model 1 for the trader group was chosen as a second-order autoregressive model based on the residual analysis for which the Yule-Walker estimation method was applied. Results show that weekly price increases generally do not affect the traders' corn purchasing behavior throughout the four different seasons. Furthermore, even when looking in more detail at the price effects during the different seasons, no significant influence on the commercial traders' corn purchasing behavior was observed during any season. This result is not unexpected as traders are responsible for supplying the manufacturing facilities year-round to keep up production.

Model 2 commercial grain traders

Model 2 for the traders was also tested with the second-order autoregressive Yule-Walker estimation. Concerning model 2, estimates revealed insignificant interaction effects for futures price changes observed for the *PreWASDE*, *WASDE*, and *PostWASDE* weeks. Results presented in table 7 reflect the estimated independent or separate effects of weekly futures price changes and *WASDE* and *PostWASDE* weeks on quantities of corn purchased by commercial traders. Results reveal that the quantities of corn purchased by commercial traders during the week before the *WASDE* release weeks were significantly positive at the 5% level, with a coefficient of 0.24 for the intercept term. This indicates that traders' weekly corn purchases increase by 0.24% during the *PreWASDE* period. Results also reveal that the corn quantities purchased by commercial traders during the week immediately following a *WASDE* release date are significantly lower compared to the *PreWASDE* period. During the *WASDE* week, commercial traders' change in weekly corn purchases declined by 0.41% compared to the *PreWASDE* weeks. The decline during the WASDE period was also significant at the 5% level. This would be consistent with the notion that increased volatility in futures prices and basis levels immediately following a report release increase the risk of buying corn during these weeks. Hence, the quantity of purchases is lower. The *PostWASDE* weeks showed insignificant changes in purchased quantities of corn by commercial traders, compared to the *PreWASDE* period.

Model 3 commercial grain traders

The final 3rd model for the commercial traders focuses on the change from the *PreWASDE* to *WASDE* weeks by isolating the *PreWASDE* weeks in the data and adding the

Mean Surprise variable with its respective interaction terms to capture the surprise factor of the WASDE report compared to private forecasts. In this case, a regular OLS model was used, and the results can be found in table 8. The results show that all variables and interaction terms are insignificant at the 5% level. At the 10% level of significance only the winter period compared to the fall intercept showed a decline of 0,46% in the purchases of commercial grain traders. Neither the change in settle price nor the mean surprise effect influenced the purchasing behavior of the commercial corn traders in our sample.

Corn Producers

The second part of the results section shows the same models as for the commercial grain traders, but this time, the models seek to reveal insights into the decision-making of the corn producers. Again, the corn producers are the counterpart to the commercial grain traders, and the producers make selling decisions rather than purchasing decisions.

Model 1 corn producers

Model 1 for the corn producers was tested with a second-order autoregressive Yule-Walker model to test whether any seasons or changes in the weekly average settle price would influence the selling behavior of corn producers. Table 9 displays the results of model 1 for the corn producers, showing that the coefficients for the different seasons show no significance per se, except summer. Sales during the summer season are slightly higher by 0.19%, significant at the 5% level, over the fall which may be caused by producers wanting to clear out their storage for the new coming harvest.

The coefficient for the *Diff Settle* is insignificant, but during the winter season, the *Diff Settle* is highly significant (the interaction term between price change and winter), at the 1% level. An increase in *Diff Settle* during the winter season of 1% would cause corn producers to increase their sales by 12.64%. Corn sales during the winter months are a popular time for farmers to sell their physical commodities. The prices are typically higher than during the harvest period, and the farmers can profit from storing their grain. However, storing agricultural products carries risk, and producers need new cash to prepare for the new planting season, which means producers don’t want to hold on to their products for too long.

Model 2 corn producers

The 2nd model for the producers focuses on the selling behavior during the different week types, the *PreWASDE*, *WASDE*, and *PostWASDE* weeks. For this model, a second-order autoregressive Yule-Walker was utilized. The results presented in Table 10 show no significant coefficients for *Diff Settle* at the 5% level. However, at the 10% level of significance, the Diff Settle during the WASDE weeks positively influences the quantities sold by corn producers. In this case, a 1% increase in *Diff Settle* would cause the sales by producers during the *WASDE* week to increase by 5.91%. This indicates that producers likely pay attention to WASDE reports as they are aware that prices can move dramatically in the wake of WASDE releases. During times when prices increase following a report producers may use this as a signal to sell their corn. This result is consistent with the notion that producers at least in part base their marketing/selling decisions on momentum trading or price rally trading.

Model 3 corn producers

Lastly, model 3 for the producers was estimated using the OLS method, and the results are displayed in table 11. Model 3 focuses on the *WASDE* weeks of the data set and contains the *Mean Surprise* effect in addition to the *Diff Settle*. The coefficient for the *Mean Surprise* during the spring season indicates producers significantly increase sales increase during the *WASDE* weeks in the Spring season. For example, a 1% increase in the *Mean Surpris*e would mean an increase of 16.8% of corn sales by producers during the spring season for the weeks that start with a WASDE release compared to the *PreWASDE* weeks, significant at the 1% level. This strong reaction to the *Mean Surprise* confirms the finding of prior studies (e.g., McKenzie, 2008; Lehecka, 2014) that USDA forecasts are valuable market information despite other private reports and other available data. However, the direction of the impact is counterintuitive. Larger surprises should be associated with larger than expected supply and lower prices. Therefore, it is not intuitive why producers would increase sales based on such a signal? During the spring season, the *Diff Settle* also shows significance at the 10% level, where a 1% increase in *Diff Settle* would cause the spring sales of corn by producers during *WASDE* weeks to increase by 25.86%. Compared to model 1, where producers sold more corn during the winter while disregarding the week types and the *Mean Surprise*, producers may be more sensitive to economic information during the spring. Producers may have sold a rather larger share of their production before the spring season. They may try to time the market to their best advantage during the spring to sell their remaining stock of corn, hence the sensitivity to the WASDE report release and the *WASDE Surprise* information.

Discussion

From the results, we can see that the seasons do not majorly influence the corn purchasing behavior of commercial grain traders. Also, WASDE reports were not of significant importance to the corn purchasing behavior of those commercial traders. We found grain traders' purchases were not affected by the change from the *PreWASDE* week to the *WASDE* week or by the surprise factor of the WASDE estimates. This result is perhaps not surprising given that purchases must be made year-round and that the traders are not pricing sensitive but instead make their buying decisions based on the basis and the need to keep feed mills continually stocked.

On the other hand, the producers seem to adjust their corn sales throughout the year, as can be observed with sold quantities of corn differing for the seasons. For example, during the winter, the *Diff Settle* led to a steep increase in sales by producers, but during the summer, the price was not the driving factor for a slight increase in sales compared to the harvest period. When looking particularly at the WASDE publication and the change from the *PreWASDE* to the *WASDE* week, the results show that the *Mean Surprise* in the spring and had a considerable influence on the producer's sales. While we do not know to which extent the producers have access to private corn supply and demand forecasts, the information content of the WASDE is valuable information that the producers use to make selling decisions, at least during the spring season. Mattos and Zinn (2016) found that grain selling producers sell more when the current market price is above the producers' reference price. The producers in our study then to react to the *Mean Surprise* in the spring, when corn process are traditionally rising. A higher Mean Surprise would mean a higher stocks-to-use ratio and lower prices, making this observation unexpected. Higher sales at prices that are expected to drop is counterintuitive to the theory of producers using the *Mean Surprise* as a reference for their sales decisions. On the other hand it could be a sign of conservatism bias (Barberis, Shleifer, & Vishny, 1998) where the producers are slow to update their beliefs to new evidence.

We also observe from the results that the commercial grain traders adjust their purchasing behavior of corn during the different week types. Meaning the release of the WASDE report triggers a response in the purchasing behavior of corn. Most corn is purchased in *PreWASDE* week, meaning the seven-day period before the WASDE release. During the week starting with the publication of a WASDE report by the USDA, the purchases of corn initiated by the commercial grain traders of our sample declined significantly compared to the *PreWASDE* week.

Since the purchased corn quantities by traders do not significantly react to the WASDE releases, but the traders purchase higher quantities in the *PreWASDE* weeks, it can be assumed that the traders value the information provided in the private forecasts and adjust their trading. While the role of private forecasts is somewhat controversial with the literature concerning the validity of public spending for the WASDE report, Englkraut et al. (2003) credited the private forecasts with providing a window of opportunity for the paying subscribers of such paid supply and demand predictions. While the reports were not claimed to be superior in their forecasts, they give the paying subscribers an advantage in acting on the market before the public does. Similarly, McKenzie (2008) also credited the private forecasts with being able to provide a good outlook on market expectations.

Limitations

The sample of commercial corn traders and producers has a rather limited number of observations on either side. It would be difficult to draw definitive conclusions for the whole population from our regionally limited sample of buyers and sellers that are all associated to one grain purchasing company. Further, the commercial grain traders in this sample are somewhat limited in their purchasing decisions, as they have to ensure the prevention of a production stop at the feed mills that they purchase inputs for. Ideally the tests of this study could be conducted with a broader sample of especially commercial grain traders.

Conclusion

Our producer results confirm that WASDE reports provide useful public information that impacts prices, beyond that contained in private reports. From a public policy standpoint adequate funding to produce these reports is essential for the U.S. grain marketing system to function effectively and efficiently allocate resources.

From our limited sample, it is challenging to measure the exact extent of the WASDE report and its forecasts as a reference point for traders and producers of corn. However, there is evidence of its importance as a public good. While only the producers react to the *Mean Surprise* directly, the WADE report publication schedule is essential to traders, who may indirectly be affected by the WASDE. In addition, the WASDE is a reference point for the publishers of private information who are motivated by profit in providing a better and faster forecast than the USDA.

References

Adjemian, M. K. (2012). Quantifying the WASDE announcement effect. *American Journal of Agricultural Economics*, *94*(1), 238-256.

Adjemian, M. K., Johansson, R., McKenzie, A., & Thomsen, M. (2018). Was the missing 2013 WASDE missed? *Applied Economic Perspectives and Policy*, *40*(4), 653-671.

Barberis, N., Shleifer, A., & Vishny, R. W. (2005). *A model of investor sentiment* (pp. 423-459). Princeton University Press.

Baucells, M., Weber, M., & Welfens, F. (2011). Reference-point formation and updating. Management Science, 57(3), 506-519. https://doi.org/10.1287/mnsc.1100.1286

Bianchi, R. J., Drew, M. E., & Fan, J. H. (2016). Commodities momentum: A behavioral perspective. *Journal of Banking & Finance*, *72*, 133-150.

CME Group. (n.d.). *Trading Hours: Futures & Options*. Retrieved October 19, 2021, from https://www.cmegroup.com/trading-hours.html

Egelkraut, T. M., Garcia, P., Irwin, S. H., & Good, D. L. (2003). An evaluation of crop forecast accuracy for corn and soybeans: USDA and private information agencies. *Journal of Agricultural and Applied Economics*, *35*(1379-2016-113519), 79-95.

Epley, N., (2013). Anchoring. In Pahler, H. (Ed.), *Encyclopedia of the Mind*

Gallant, A. R., and Goebel, J. J. (1976). Nonlinear Regression with Autoregressive Errors. *Journal of the American Statistical Association*, 71:961–967.

Isengildina-Massa, O., Irwin, S. H., Good, D. L., & Gomez, J. K. (2008). The impact of situation and outlook information in corn and soybean futures markets: Evidence from WASDE reports. *Journal of Agricultural and Applied Economics*, *40*(1379-2016-112694), 89-103.

Isengildina-Massa, O., Karali, B., & Irwin, S. H. (2013). When do the USDA forecasters make mistakes? *Applied Economics*, *45*(36), 5086-5103.

Isengildina-Massa, O., Karali, B., & Irwin, S. H. (2020). Can private forecasters beat the USDA? analysis of relative accuracy of crop acreage and production forecasts. *Journal of Agricultural and Applied Economics*, *52*(4), 545-561.

Just, R. E. (1983). The impact of less data on the agricultural economy and society. *American Journal of Agricultural Economics*, *65*(5), 872-881.

Lehecka, G. V. (2014). The value of USDA crop progress and condition information: Reactions of corn and soybean futures markets. *Journal of Agricultural and Resource Economics, 39*(1), 88-105. Retrieved from https://www.jstor.org/stable/44131316

Mattos, F. L., & Fryza, S. A. (2012). Marketing Contracts, Overconfidence, and Timing in the Canadian Wheat Market. *Journal of Agricultural and Resource Economics*, 469-484.

Mattos, F. L., & Zinn, J. (2016). Formation and adaption of reference price in grain marketing: An experimental study. *Agricultural Economics*, *47*, 621-632.

McKenzie, A. M. (2008). Pre-harvest price expectations for corn: The information content of USDA reports and new crop futures. *American Journal of Agricultural Economics, 90*(2), 351-366. doi:10.1111/j.1467-8276.2007.01117.x

Microsoft.Microsoft excel [computer software]

SAS Insititue Inc. (2019). Sas [computer software]. Cary, NC, USA

SAS Institute Inc. 2015. SAS/ETS® 14.1 *User's Guide*. Cary, NC: SAS Institute Inc.

Tonsor, G. T. (2018). Producer decision making under uncertainty: Role of past experiences and question framing. American Journal of Agricultural Economics, 100(4), 1120-1135. https://doi.org/10.1093/ajae/aay034

U.S. Department of Agriculture. (2013, October). *USDA announces cancellation and postponement of selected reports impacted by the lapse in federal funding*. Retrieved from https://www.usda.gov/media/press-releases/2013/10/17/usda-announces-cancellation-and-postponement-selected-reports

U.S. Department of Agriculture. (2012, September). *USDA announces change in release time of key statistical reports beginning in january 2013*. Retrieved from https://www.usda.gov/media/press-releases/2012/09/19/usda-announces-change-release-time-key-statistical-reports

Wenner, L. M. (2015). Expected prices as reference points—Theory and experiments. European Economic Review, 75, 60-79. https://doi.org/10.1016/j.euroecorev.2015.01.001

Xie, R., Isengildina-Massa, O., Dwyer, G. P., & Sharp, J. L. (2016). The impact of public and semi-public information on cotton futures market. Applied Economics, 48(36), 3416-3431.

Tables

Table 1

Descriptive statistics for model 1

	N	Variable	Mean	Std Dev	Min	Max
Traders	231	Sum QTY (Bu)	3,670,084	2,580,575	290,000	19,613,100
		Mean Settle (USD per bu)	4.75	1.44	3.23	8.17
Producers	231	Sum QTY (Bu)	423,447	395,958	28,900	2,698,860
		Mean Settle (USD per bu)	4.75	1.44	3.23	8.17

Table 2

Descriptive statistics for model 2 for the traders

Period	N	Variable	Mean	Std Dev	Min	Max
PreWASDE	77	Sum QTY (Bu)	3,342,778	2,094,804	432,000	9,596,450
		Mean Settle (USD per bu)	4.74	1.46	3.29	8.17
WASDE	77	Sum QTY (Bu)	3,614,691	2,608,003	290,000	14,920,056
		Mean Settle (USD per bu)	4.74	1.43	3.28	7.91
Control	77	Sum QTY (Bu)	4,052,783	2,950,157	524,098	19,613,100
		Mean Settle (USD per bu)	4.76	1.45	3.23	8.08

Table 3

Descriptive statistics for model 2 for the producers

Period	N	Variable	Mean	Std Dev	Min	Max
PreWASDE	77	Sum QTY (Bu)	403,413	342,706	43,500	1,735,000
		Mean Settle (USD per bu)	4.74	1.46	3.29	8.17
WASDE	77	Sum QTY (Bu)	440,153	402,230	28,900	2,626,499
		Mean Settle (USD per bu)	4.74	1.43	3.28	7.91
Control	77	Sum QTY (Bu)	426,777	441,064	39,000	2,698,860
		Mean Settle (USD per bu)	4.76	1.45	3.23	8.08

Table 4

Descriptive statistics for model 3 for the traders

Season	N	Variable	Mean	Std Dev	Min	Max
Fall	20	Sum QTY (Bu)	3,802,301	2,029,660	1,057,000	8,204,600
		Mean Settle (USD per bu)	4.69	1.64	3.34	7.63
		Mean Surprise	0.03	0.07	-0.08	0.21
Winter	20	Sum QTY (Bu)	3,424,596	2,626,760	817,837	12,531,753
		Mean Settle (USD per bu)	4.67	1.37	3.48	7.26
		Mean Surprise	-0.01	0.05	-0.13	0.11
Spring	17	Sum QTY (Bu)	3,583,298	1,962,404	1,260,022	8,301,740
		Mean Settle (USD per bu)	4.91	1.31	3.64	7.11
		Mean Surprise	0	0.05	-0.1	0.11
Summer	19	Sum QTY (Bu)	3,703,953	3,675,069	290,000	14,920,056
		Mean Settle (USD per bu)	4.77	1.49	3.28	7.91
		Mean Surprise	0.02	0.07	-0.12	0.17

Table 5

Descriptive statistics for model 3 for the producers

Season	N	Variable	Mean	Std Dev	Min	Max
Fall	20	Sum QTY (Bu)	402,209	338,714	28,900	1,363,800
		Mean Settle (USD per bu)	4.69	1.64	3.34	7.63
		Mean Surprise	0.03	0.07	-0.08	0.21
Winter	20	Sum QTY (Bu)	590,166	612,788	46,500	2,626,499
		Mean Settle (USD per bu)	4.67	1.37	3.48	7.26
		Mean Surprise	-0.01	0.05	-0.13	0.11
Spring	18	Sum QTY (Bu)	360,847	267,148	54,411	1,252,600
		Mean Settle (USD per bu)	4.85	1.29	3.64	7.11
		Mean Surprise	0	0.05	-0.1	0.11
Summer	19	Sum QTY (Bu)	397,316	251,803	85,999	1,078,400
		Mean Settle (USD per bu)	4.77	1.49	3.28	7.91
		Mean Surprise	0.02	0.07	-0.12	0.17

Table 6

Parameter estimates for model 1 for traders

	Traders Model 1 second-order autocorrelation Yule-Walker method
Parameters	
Intercept	-0.01
	(0.05)
Difference Settle	-1.69
	(2.52)
Winter	0.03
	(0.08)
Difference Settle Winter	5.72
	(4.21)
Spring	-0.06
	(0.08)
Difference Settle Spring	5.88
	(4.06)
Summer	0.04
	(0.08)
Difference Settle Summer	2.59
	(2.78)
R^2	0.32
White's Test	0.87
	(0.65)
Breusch-Pagan	0.73
	(0.39)
Arima	14.67
	(-0.02)

Standard errors are shown in parentheses for parameters. P-values are shown in parentheses for diagnostic test statistics.
* Indicates reject the null hypothesis at the 10% level.
** Indicates reject the null hypothesis at the 5% level.
*** Indicates reject the null hypothesis at the 1% level.

Table 7

Parameter estimates for model 2 for traders

Traders Model 2 second-order autocorrelation Yule-Walker method	
Parameters	
Intercept	0.24 **
	(0.11)
Difference Settle	1.41
	(1.18)
WASDE	-0.41 **
	(0.18)
Difference Settle WASDE	-1.04
	(2.85)
PostWASDE	-0.30 *
	(0.18)
Difference Settle PostWASDE	-1.22
	(3.09)
R^2	0.32
White's Test	1.11
	(0.57)
Breusch-Pagan	1.09
	(0.30)
Arima	16.43
	(0.01)

Standard errors are shown in parentheses for parameters. P-values are shown in parentheses for diagnostic test statistics.
* Indicates reject the null hypothesis at the 10% level.
** Indicates reject the null hypothesis at the 5% level.
*** Indicates reject the null hypothesis at the 1% level.

Table 8

Parameter estimates for model 3 for traders

	Traders Model 3 OLS
Parameters	
Intercept	-0.05
	(0.19)
Difference Settle	5.26
	(7.76)
Mean Surprise	3.85
	(3.13)
Winter	-0.46 *
	(0.27)
Difference Settle Winter	0.02
	(11.21)
Mean Surprise Winter	-8.68
	(5.46)
Spring	0.02
	(0.26)
Difference Settle Spring	-9.97
	(12.66)
Mean Surprise Spring	-3.36
	(5.57)
Summer	-0.14
	(0.26)
Difference Settle Summer	-2.77
	(11.16)
Mean Surprise Summer	-3.8
	(4.22)
R^2	0.15
White's Test	0.62
	(0.73)
Breusch-Pagan	0.44
	(0.50)
Arima	9.69
	(0.14)

Standard errors are shown in parentheses for parameters. P-values are shown in parentheses for diagnostic test statistics.
* Indicates reject the null hypothesis at the 10% level.
** Indicates reject the null hypothesis at the 5% level.
*** Indicates reject the null hypothesis at the 1% level.

Table 9

Parameter estimates for model 1 for producers

Producers Model 1 second-order autocorrelation Yule-Walker method	
Parameters	
Intercept	-0.10
	(0.06)
Difference Settle	0.31
	(2.80)
Winter	0.15
	(0.08)
Difference Settle Winter	12.64 ***
	(4.75)
Spring	0.03
	(0.08)
Difference Settle Spring	7.21
	(4.56)
Summer	0.19 **
	(0.09)
Difference Settle Summer	2.20
	(3.10)
R^2	0.37
White's Test	2.87
	(0.24)
Breusch-Pagan	0.42
	(0.52)
Arima	11.89
	(0.06)

Standard errors are shown in parentheses for parameters. P-values are shown in parentheses for diagnostic test statistics.
* Indicates reject the null hypothesis at the 10% level.
** Indicates reject the null hypothesis at the 5% level.
*** Indicates reject the null hypothesis at the 1% level.

Table 10

Parameter estimates for model 2 for producers

	Producers Model 2 second-order autocorrelation Yule-Walker method
Parameters	
Intercept	-0.01
	(0.14)
Difference Settle	1.84
	(1.39)
WASDE	0.13
	(0.23)
Difference Settle WASDE	5.91*
	(3.36)
PostWASDE	-0.05
	(0.23)
Difference Settle PostWASDE	5.06
	(3.64)
R^2	0.34
White's Test	3.00
	(0.22)
Breusch-Pagan	2.11
	(0.15)
Arima	9.72
	(0.14)

Standard errors are shown in parentheses for parameters. P-values are shown in parentheses for diagnostic test statistics.
* Indicates reject the null hypothesis at the 10% level.
** Indicates reject the null hypothesis at the 5% level.
*** Indicates reject the null hypothesis at the 1% level.

Table 11

Parameter estimates for model 3 for producers

	Traders Model 3 OLS
Parameters	
Intercept	0.08
	(0.21)
Difference Settle	4.95
	(8.59)
Mean Surprise	-0.07
	(3.47)
Winter	-0.02
	(0.30)
Difference Settle Winter	15.39
	(12.41)
Mean Surprise Winter	-2.95
	(6.04)
Spring	0.46
	(0.29)
Difference Settle Spring	25.89 *
	(14.02)
Mean Surprise Spring	16.80 ***
	(6.16)
Summer	-0.18
	(0.29)
Difference Settle Summer	1.20
	(12.35)
Mean Surprise Summer	-2.82
	(4.67)
R^2	0.33
White's Test	2.18
	(0.34)
Breusch-Pagan	0.01
	(0.92)
Arima	9.97
	(0.13)

Standard errors are shown in parentheses for parameters. P-values are shown in parentheses for diagnostic test statistics.
* Indicates reject the null hypothesis at the 10% level.
** Indicates reject the null hypothesis at the 5% level.
*** Indicates reject the null hypothesis at the 1% level.

Conclusion

This research focuses on gaining new insights into the decision-making of agricultural commodity markets to improve training for private and commercial parties and guide policy decisions on public spending for the provision of market information. In the first article, we look at the trading performance of a set of commercial grain traders. We are particularly interested if individual traders possess the ability to consistently outperform the soybean or corn cash markets by using forward contracts. The traders differ from typically investigated speculative traders in that they use forward trades as a price risk management tool to buy input commodities for feed mills. For this first article, the data was set up in yearly and quarterly periods. Winner and loser rank tests and top and bottom performance tests were used to rank and evaluate trader performance. The results of the data analysis do not show proof that individual traders are able to persistently outperform the market. In the short run, some traders may achieve repeated top rankings, and the winners are able to outperform the losers in terms of achieved margin in consecutive quarterly periods. When comparing the achieved margins of the top trader to the bottom trader in consecutive time periods, their performance differs significantly, with the previous trader being significantly more successful in achieving a greater margin again in the later time period than the bottom trader of the earlier period. These results cannot be confirmed for the yearly data where trader ranking is random. The results show that some traders can profit from superior price movement forecasts in the short run, but superior abilities are not observed in the long run. Further, the first paper also investigates the role of gender and the experience level of traders as a determinant of trading performance in terms of margin. Both tests for gender and experience were conducted for a single trading year, limiting their validity. Surprisingly we found that inexperienced traders were able to achieve better margins than experienced trades. A possible explanation would be increased risk-taking by more experienced traders that may have led to the unexpected result. As for gender, the women were able to outperform their male colleagues during the 2017 season, but we could not confirm this finding when considering all trades over a 6-year period. Considering the results of the several different tests we can confirm that the performance of commercial grain traders is often unexpected and difficult to explain. Understanding decision-making and trader behavior require further research to make markets more efficient.

The second paper applies a panel data analysis approach to understand how the previous performance of traders influences their future trading behavior. The future trading behavior is measured in the quantities bought with forward contracts and the length of the contracts that traders engage in. Both the forward contracted quantities and contract length increase in risk when quantities are larger, or contracts get longer. The panel data analysis shows no evidence of the traders' past performance, as measured in the average margin of the previous trading period, on either the quantities purchased with forward contracts or the contract length. The only impactful factor for trader performance was the average gateway basis when measuring contract length. Traders react, as expected, negatively when the derived gateway basis is high at the point of contract agreement and the contract length in days becomes shorter.

In conclusion, the lack of reaction of the traders to their past performance and the importance of the derived gateway prices may point to the importance of food-producing companies to reduce price risk rather than their traders achieving the highest margin possible in a risky environment. Proven skill to operate in agricultural commodity markets with tight margins is the use of basis trading. The results show that risk mitigation is of great interest and that basis trading techniques can help make better purchasing decisions. Therefore, efforts to promote basis trading should not only be limited to private companies but also agricultural producers, who can also benefit from a deeper understanding of commodity markets to improve their selling decisions and, ultimately, their profitability.

Finally, in the third paper, the role of public information on trader and seller behavior in agricultural commodity markets is studied. Three models were tested each for the purchasing behavior of commercial grain traders and the selling behavior of corn producers. Regression analysis methods suitable for time series data were employed to understand the behavior of the two different groups of market participants. The studied public information in this paper is the World Agricultural Supply and Demand (WASDE) released monthly by the USDA. We are particularly interested in the stocks-to-use ratio published in the report in relation to stocks-to-use ratios published by private agencies shortly before the release of the public WASDE. The data was categorized into three weekly periods consisting of the week prior to the WASDE release, the second week starting with the WASDE release, and the week after the WASDE was released. Besides the categorization into weeks, the sales and purchase datasets were also categorized by season. In the first test, the weekly differences in quantities sold and purchased were compared by season. Both parties, the traders and the producers showed different trading activities throughout the seasons. First, the traders are not changing the amount of bushels they purchase throughout the seasons and neither did the difference in settle price (*Diff Settle*) or the *Diff Settle* and season interaction terms impact the purchasing behavior of commercial corn traders. On the other side, the corn producers significantly increased their corn sales based on the *Diff Settle* during the winter period. Corn producers also sold larger quantities during the summer season, without the *Diff Settle* being the driving factor during the summer season. The second test investigated the role of the weekly periods on the selling and purchasing behavior of corn.

The commercial grain traders purchased more grain during the 7-day period prior to the WASDE report publication, *PreWASDE*. Their purchased amount of bushels purchased dropped significantly for the 7-day period starting with the WASDE publication and also was lower during the 7-day control period post *WASDE* release week. The interactions of the *Diff Settle* did not seem to impact the purchased amounts of the commercial grain traders. The second test did not yield highly significant results for the corn producers, who showed a slight increase of corn sales during the *WASDE* week when the *Diff Settle* rose. The final third test of this paper took a closer look at the change form the *PreWASDE* weeks to the *WASDE* weeks with consideration of the *Diff Settle* and the *Mean Surprise* indicating the surprise information in the WASDE of stocks-to-use forecasts after the publication of private reports also estimating the stocks-to-use ratio. Again like for the first test that tested the seasonal changes in purchases only, the more detailed look at the change from the *PreWASDE* to *WASDE* week did not show much significance for the commercial corn traders. The corn producers on the other hand showed a strong reaction to the *Mean Surprise* during the spring season when the producers unexpectedly increased their sales as a reaction to an increase in the *Mean Surprise*, which should indicate a higher supply and lower expected prices for corn. If the producers used the Mean Surprise as a reference point their decisions, we would expect different behavior.

Our results show that the behavior of the different market participants in the agricultural markets is still not fully understood and that further research would be valuable to inform policy makers of suitable tools to support all market participants in the provision of information and training, while allocating financial resources in the most efficient way possible.

Appendix

Research Approval Letter

UNIVERSITY OF ARKANSAS

To:	Marei Undine Houpert MULN 220Q
From:	Douglas James Adams, Chair IRB Committee
Date:	09/28/2018
Action:	**Expedited Approval**
Action Date:	09/24/2018
Protocol #:	1809144820
Study Title:	Decision making under uncertainty among agricultural commodity traders
Expiration Date:	09/23/2019
Last Approval Date:	

The above-referenced protocol has been approved following expedited review by the IRB Committee that oversees research with human subjects.

If the research involves collaboration with another institution then the research cannot commence until the Committee receives written notification of approval from the collaborating institution's IRB.

It is the Principal Investigator's responsibility to obtain review and continued approval before the expiration date.

Protocols are approved for a maximum period of one year. You may not continue any research activity beyond the expiration date without Committee approval. Please submit continuation requests early enough to allow sufficient time for review. Failure to receive approval for continuation before the expiration date will result in the automatic suspension of the approval of this protocol. Information collected following suspension is unapproved research and cannot be reported or published as research data. If you do not wish continued approval, please notify the Committee of the study closure.

Adverse Events: Any serious or unexpected adverse event must be reported to the IRB Committee within 48 hours. All other adverse events should be reported within 10 working days.

Amendments: If you wish to change any aspect of this study, such as the procedures, the consent forms, study personnel, or number of participants, please submit an amendment to the IRB. All changes must be approved by the IRB Committee before they can be initiated.

You must maintain a research file for at least 3 years after completion of the study. This file should include all correspondence with the IRB Committee, original signed consent forms, and study data.

cc: Andrew Malcolm McKenzie, Investigator
Rodolfo M Nayga Jr., Key Personnel