ECONOMIC THRESHOLDS OF WHEAT STREAK MOSAIC, FEEDLOT COST OF GAIN PREDICTION, AND JOINT ADOPTION OF COW-CALF PRODUCTION PRACTICES

By

BRIAN P. MULENGA

Bachelor of Science in Agricultural Economics
University of Zambia
Lusaka, Zambia
2007

Master of Science in Agricultural and Applied Economics
University of Malawi
Lilongwe, Malawi
2011

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 2019

ECONOMIC THRESHOLDS OF WHEAT STREAK MOSAIC, FEEDLOT COST OF GAIN PREDICTION, AND JOINT ADOPTION OF COW-CALF PRODUCTION PRACTICES

Dissertation Approved:

Dr. Kellie Raper

Dissertation Adviser

Dr. Derrell Peel

Dr. Wade Brorsen

Dr. David Lalman

.

ACKNOWLEDGEMENTS

I thank God Almighty for giving me the strength and motivation to be able to complete my doctorate program. Numerous people contributed to this accomplishment and I wish to express my heartfelt gratitude to them.

To start with, I wish to express my deepest and sincere gratitude to my major advisor, Dr. Kellie Raper for her kindness, enduring patience, unwavering support and encouragement. She made me believe in myself more than I did before I started working with her. Working with her helped me appreciate the interface between academic research and extension. This work would not have been possible without her advice and guidance. I also wish to thank my committee members, Dr. Derrell Peel, Dr. Wade Brorsen, and Dr. Lalman for their invaluable feedback, especially at the time I presented my proposal and during my dissertation defense. I would be failing in my acknowledgements if I did not thank Dr. Francis Epplin and Dr. Chanjin Chung for the advisory role they played before I was assigned a major advisor. Their commitment to help me settle in the program despite not being my major advisors is priceless. Let me also say enormous thanks to Ms. Anna Whitney, the graduate student specialist, for her efficient handling of all my administrative matters, and ensuring that my paperwork was always up to date and in order.

To my late mom (Christine Mulenga), late dad (Widson Ng'uni), and late grandmother (Evelyn Nyondo), words can never describe how grateful I am for preparing me for the joys and challenges of life. I try my best to always carry with me the virtues of life you taught me including honesty, hard work, and respect among many others. Your faith in my abilities has kept me going thus far.

My special thanks go to my wife (Bernadette Chimai), who is my best friend and happens to be my classmate, not only in the Ph.D. program but also way back in college. She has seen me at my strongest and weakest points and everything in between, and still remains supportive. To our two beautiful daughters Mphatso and Mwiza, thank you for your patience and understanding. There were countless times I could not take you out to play because I was either studying or working on my research, and though young, you understood. I hope to make it up to you some day.

Last but not least, I would like to thank my fellow graduate students who in one way or the other contributed to my successful completion of the program. May God Almighty bless you.

Name: BRIAN P. MULENGA

Date of Degree: MAY, 2019

Title of Study: ECONOMIC THRESHOLDS OF WHEAT STREAK MOSAIC, FEEDLOT COST OF GAIN PREDICTION, AND JOINT ADOPTION OF COW-CALF PRODUCTION PRACTICES

Major Field: AGRICULTURAL ECONOMICS

Abstract:

Wheat streak mosaic (WSM), caused by *Wheat streak mosaic virus*, which is transmitted by the wheat curl mite (*Aceria tosichella* Keifer), is the most widespread and economically important virus disease affecting winter wheat in the Great Plains of the United States. Because there is no curative treatment, the disease can lead to significant yield loss, rendering continuation of mid-season input application uneconomical. This dissertation determines three economic thresholds for WSM beyond which further input applications become uneconomical. Results show varying thresholds depending on the date of disease severity assessments. Results indicate potential to save resources by discontinuing mid-season input applications and introducing cattle for grazing, in about 14% of the sampled plots.

Feedlots use cost of gain (COG) to evaluate the tradeoff between purchasing heavier feeders or lighter feeders. Typically, ex-post COG for the feeding period just finished is used as a naïve projection for future closeout. However, such a naïve estimate may mask the effect of varying corn price and seasonality of cattle feeding efficiency on COG. This dissertation constructs an ex-ante COG prediction model to help facilitate more accurate estimation of expected COG, and thus corresponding feedlot purchase breakeven price, for lighter animals relative to heavier animals (800-850 pounds). Results show that the constructed model predicts COG more accurately than the naïve model. Results also show small reduction in price spreads between feeder market price and breakeven price for lighter weights when the predicted COG is used in calculating the breakeven price.

Existing studies on adoption of cow-calf management practices tend to treat practices individually and by implication ignore the possibility that some practices are more likely to be jointly adopted. This dissertation uses market basket analysis to bundle practices based on the likelihood of joint adoption. Results show that dehorning plus polled genetics (horn management), deworming, and castration are the top three most widely adopted practices and are more likely to be jointly adopted in varying combinations with other practices. Results indicate higher conditional likelihood of vaccination if both feed bunks and 45-day weaning are adopted in addition to the top three practices.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

This dissertation is comprised of three essays focused on the economics of wheat and beef cattle production. The first essay analyzes the economic thresholds of wheat streak mosaic (WSM), using field experiment data from the Southern Great Plains region. The second essay uses monthly data from the Kansas feedlot performance and feed cost summary spanning January 1992 to July 2017, to model feedlot cost of gain prediction and examine implications for breakeven analysis. The third essay uses producer level survey data from Oklahoma cow/calf producers to model the joint adoption of value-added beef production practices.

Essay I (Chapter II) focuses on estimating the economic thresholds of WSM and how such information can help wheat producers in WSM affected areas make profit maximization decisions when their crop is infected. *Wheat streak mosaic virus* (WSMV) causes wheat streak mosaic (WSM) and is transmitted by the wheat curl mite (*Aceria tosichella*). The disease may be confined to a specific part of the field or can spread to the entire field and to adjacent fields. Infection generally occurs. In the fall if green vegetative plants such as volunteer wheat, native green grasses, and even corn, infested with the virus-carrying wheat curl mites are nearby when wheat seedlings emerge. The diseases has been found to spread from an infected field to a healthy field a distance of 1.4 miles. Although infections can occur both in the fall, soon after plants emerge, and in the spring, after the crop comes out of dormancy, the disease may not be noticed until temperatures begin to warm in late spring.

Common symptoms of the disease include chlorosis, streaking, and mosaic, and also stunting when plants are infected at an early stage. Infection may occur at any stage of development but if infection occurs during the early stages of crop development, the effects on crop growth and yield are more severe. As there is no curative treatment, there is little a farmer can do once the crop is infected.

Previous research has mainly focused on the effect of the disease on yield and yield determinants such as tillering, shoot weight, and plant height, which have been found to be significantly reduced by WSMV infection. With a few exceptions, most studies use descriptive analysis to estimate the correlation between disease severity and yield. However, information on the effects on WSM on final yields may not be very useful to a producer. Rather, what would be more useful to a producer is information that will help signal whether continuing with input application beyond a certain infection severity is economical. This is because a high infection severity at a given point in the season, could substantially reduce final yields such that continuing with mid-season input applications becomes uneconomical. Under such circumstances, a producer can improve returns and save resources by ceasing input applications and graze-out the wheat. These studies provided important insights into the potential effect of WSM on yield. The current study presents a more nuanced yield response estimate, which is used to estimate an economic infection threshold, beyond which it becomes uneconomical for a producer to continue with input applications. Determining the economic threshold for WSM early in the growing season so farmers could discontinue input applications, has potential to save resources and reduce costs.

With this goal in mind, the study used field experiment data and input costs information from a wheat grower to help address the following two objectives; 1) determine wheat grain yield response to WSM severity as estimated by remote sensing; and 2) determine the disease severity threshold, beyond which it is uneconomical to continue with crop inputs.

The second essay (Chapter III) is concerned with modeling expected feedlot cost of gain, an important cost component of total feeding costs. The decision by feedlots regarding the size of feeder cattle to place on feed depends, in part, on the expect closeout cost of gain. Feedlots use cost of gain to evaluate the tradeoff between purchasing lighter or heavier feeder cattle. Generally, this tradeoff is between two decisions; 1) purchase more pounds via heavier feeders and spend less on feed to bring the cattle to slaughter weight; or 2) purchase lighter feeders and put more weight on in the feedlot, which implies higher feeding costs to bring the animal to slaughter weight. Hence, cost of gain can be used to estimate a feedlot purchase breakeven price for lighter weight feeders compared to heavy feeders. Typically, monthly data on ex-post COG estimates for current feedlot closeouts (i.e., the feeding period just finished) is used as a naïve estimate for future COG. However, such a naïve estimate is not able to capture the strong relationship between COG and corn price, and the seasonal variation of this relationship.

Given that corn is an important ingredient in feedlot rations, its price variation over-time is expected to affect COG. Further, there is seasonal variation in the relationship between feeder prices and feedlot breakeven price by weight, which may not be captured by the naïve model. Thus, it may be possible to improve understanding of this relationship by constructing a better ex-ante COG prediction model. An improved model could facilitate more accurate estimation of expected COG, and thus corresponding feedlot purchase breakeven price, for lighter animals relative to heavier animals. Most existing studies use a static simple regression of profit or COG on corn futures price and feeder cattle price, among other variables, to predict cost of gain, hence unable to model for seasonal variations in the COG-corn price relationship. Other studies model COG-corn price relationship as an autoregressive process using monthly time series data aggregated at the national level. To the extent that such studies are able to control for seasonal variation, they are able to capture the dynamic relationship between COG and corn price. However, the use of nationally aggregated data will not capture regional variations in terms of feeding dynamics, availability of corn substitutes, and market conditions. Further, existing studies use data that are now dated, thereby warranting a need to extend and update the models.

The analysis in chapter III builds on previous studies by using updated regional data, to construct a COG prediction model that incorporates the dynamics of the COG - corn price relationship and seasonality to more accurately predict expected COG and thus feedlot breakeven purchase prices. Findings of this analysis is also expected to aid evaluate the extent to which the difference between the current ex-post COG (naïve model) and ex-ante COG estimates explain the observed seasonal variations in the relationship between feeder market prices and feedlot breakeven price by weight.

Against this backdrop, the second essay has the following two objectives: 1) to construct a feeder cattle cost of gain prediction model; and 2) to estimate feedlot breakeven price across weight groups based on predicted cost of gain, and compare it with feeder market price.

The third essay (Chapter IV) focuses on the cow-calf segment of the beef value chain, specifically looking at producer adoption of value added beef production practices. Research pertaining to value addition of recommended cow-calf production practices has shown improved producer returns from adopting the value added production practices, which include castration, dehorning, deworming, 45-day weaning before sell to backgrounders, and vaccination for respiratory infections, among others. Despite the documented benefits of adopting these practices, their adoption among Oklahoma producers remains lower than expected. Understanding the producer adoption decision is critical in helping design effective extension programs aimed at encouraging producer adoption of practices deemed important by research. Existing studies regarding practice adoption has tended to treat practices independently of each other, when in reality some practices, such as castration and dehorning are more likely to be adopted together, and thus correlated. This implicit imposition of independence across practices could lead to inaccurate estimates of adoption likelihood since resulting probabilities are unconditional rather than conditional on correlation practices.

The current study models for correlation among practices by creating practice bundles based on the likelihood of joint adoption, thus relaxing the independence restrictions imposed by other studies. Further, given that data used in most studies is becoming dated, a study using recent data to update understanding of the subject is warranted. The current study uses data from a 2018 survey of Oklahoma cow-calf producers, and thus provides an updated understanding of producer practice adoption. This information could be useful in improving targeting and design of extension programs. To this end, the objectives of the third essay are as follow: 1) identify and construct clusters (bundles) of practices likely to be jointly adopted; and 2) estimate the conditional probability associated with adopting an additional practice given adoption of a bundle of practices.

CHAPTER II

ECONOMIC THRESHOLDS OF WHEAT STREAK MOSAIC IN THE TEXAS HIGH PLAINS

Introduction

The U.S. southwestern Great Plains is a major wheat production region. In addition to the millions of acres of dryland production, several million acres are produced under center pivot irrigation, which increases yields but also increases production costs. The majority of wheat grown in this region also is dual purpose, i.e., grown as winter forage for cattle grazing, and as a grain crop. Dual purpose wheat is typically planted in late August to early September, significantly earlier than wheat planted strictly for grain production, which is usually planted from late September to early November. Cattle are placed on dual purpose wheat fields in November and are allowed to graze until early March. At this time a decision is made to remove the cattle and continue the crop for grain production or to graze it out. The choice to continue or graze out is typically based on the farmer's estimate of the field's yield potential for grain production, and on commodity prices for wheat and beef. Although early planting of dual purpose wheat is necessary to maximize forage production, it also results in a longer exposure of the crop to a variety of insect pests and diseases, before the onset of winter dormancy. Among these biotic threats, wheat streak mosaic and other mite-vectored virus diseases are among the most common and economically significant (Burrows et al. 2009; Velandia et al. 2010; Workneh et al. 2017).

Wheat streak mosaic virus (WSMV) that causes wheat streak mosaic (WSM) is transmitted by the wheat curl mite (*Aceria tosichella*). The disease may be confined to a specific part of the field or can spread to the entire field and to adjacent fields. Infection generally occurs in the fall if green vegetative plants such as volunteer wheat, native green grasses, and even corn, infested with the virus-carrying wheat curl mites are nearby when wheat seedlings emerge. Volunteer wheat provides a convenient green bridge (plants growing between the harvest of last year's wheat crop and emergence of the new wheat crop) for the mites, and thus WSMV disease outbreaks (e.g., Michael and William 1993; McMullen and Waldstein 2010; Price 2015). Eliminating green bridges destroys the mites' food source and thus the mites do not live to infest subsequent crops.

When volunteer wheat is left growing late in the summer, after harvest, mites move from volunteer wheat to newly planted fields of winter wheat, completing the green bridge, with newly emerged wheat plants now hosting viruliferous mites. When conditions for disease development are conducive, i.e., warm temperatures coupled with wind, during the fall, wheat curl mites transmit the virus from infected plants (Michael and William 1993; Price 2015) to other parts of the field and to other fields in the vicinity. The disease has been found to spread from an infested field to a healthy field, a distance of 1.4 miles (McMullen and Waldstein 2010). Although infections can occur both in the fall, soon after plant emergence, and in the spring, after the crop comes out of dormancy, the disease may not be noticed until temperatures begin to warm in late spring.

Previous research found that the severity of WSM can be quantified by remote sensing, based on tissue reflectance, when the crop is at Feekes growth stage of 5 to 9 (usually about mid-April in the Southern Plains), (Mirik et al. 2013; Workneh et al. 2010, 2017). Common symptoms of the disease include chlorosis, streaking, and mosaic, and also stunting when plants are infected at an early stage. Infection may occur at any stage of development but if infection occurs during the early stages of crop development, the effects on crop growth and yield are more severe

(Workneh et al. 2009). As there is no curative treatment, there is little a farmer can do once the crop is infected. Therefore, eradication of green bridge hosts, planting resistant or tolerant varieties and avoiding early planting dates are the only means of reducing risk of loss from this disease (Byamukama et al. 2014).

Given the economic significance of WSM to wheat production and profitability, it should be beneficial to model and quantify wheat yield response to varying levels of WSM infection. In addition, it is also important to estimate the level of WSM severity which would reduce yield enough that discontinuing application of inputs to the crop and grazing it out as pasture would be a more economical option. Determining this level of infection is important as it would equip farmers with information to make informed decisions soon enough in the season to enable reductions in input applications and associated expenses.

Literature on the relationship between WSM and yield is mainly based on descriptive analysis. A few exceptions are Workneh et al. (2017), Almas et al. (2016), Pradhan, et al. (2015), and Byamukama et al. (2014), who used regression analysis to model this relationship. Byamukama et al. (2014) focused on the effect of the disease on yield determinants such as tillering, shoot weight, and plant height, which they found to be significantly reduced by WSMV infection. Workneh et al. (2017) and Almas et al. (2016) estimated regressions and found an exponential relationship between WSM and wheat yield. Almas et al. (2016) treated WSM reflectance readings (sensing variable) as a discrete variable, rather than continuous, which could potentially lose information.

These studies provided important insights into the potential effect of WSM on yield. The current study builds on these analyses and presents a more nuanced yield response estimate, which is used to estimate an economic infection threshold. Most existing literature on disease thresholds focuses on optimal timing of treatment for crop disease control (Kuosmanen 2006; Mbah et al. 2010). However, since there is no curative treatment for WSM, determining an economic threshold for WSM early in the growing season so farmers could discontinue input applications, has potential to save resources and reduce costs. With this goal in mind, the objectives of the current study were twofold; 1) determine wheat grain yield response to WSM severity as estimated by remote sensing; and 2) determine the disease severity threshold, beyond which it is uneconomical to continue with crop inputs.

Theory

Previous research has found that reflectance measurements at 555 nm are positively correlated with severity of WSM symptoms, which are negatively correlated with final grain yields (e.g., Workneh et al., 2009; Almas et al. 2016). At Feekes growth stage of 5 to 6, reflectance readings from healthy plants generally are around 5 at maximum depending on the cultivar: light-colored healthy cultivars may give higher readings than dark-green ones. Almas et al. (2016) used 4 as a maximum value for healthy, WSM-free plants. A field with a high incidence of WSM may go undetected by a grower until late in the growing season, at which point the yield potential of severely infected plants will be significantly reduced relative to uninfected plants. However, because incidence and severity of WSM in a field is progressive over time (Workneh et al. 2009), parts of a field may be severely infected while other parts of the field may be healthy. Such a scenario makes it extremely difficult for a grower to know whether additional crop inputs for the entire field are worth the expense. However, with the advent of precision agriculture and site-specific management options, growers can now apply different management strategies to different parts of a field and focus on those parts of the field with the highest yield potential.

WSM severity estimates between late March and early May can be made using remote sensing technologies to determine the relationship between disease severity, at a given time, and yield potential of the infected part of the field. Based on this information, a farmer could determine whether to continue to invest inputs in the crop or to abandon the entire crop, or at least the part of the crop with low yield potential. To salvage some value from the abandoned crop or part of the field, a farmer could harvest the wheat for hay or bring in cattle and graze it out. Our research seeks to determine optimal post remote sensing strategies, based on the sensor reading level, i.e., disease severity. After scanning the crop with remote sensing, the expected optimal strategy may be to: (1) Continue with input applications and harvest; (2) Abandon the field, or at least the portion with the lowest yield potential, discontinue input applications and graze out or harvest for hay; and (3) discontinue input applications, but let the crop mature and harvest (expected yield is greater than harvest cost)[1].

Farmers are faced with deciding whether to continue applying in-season management inputs, such as fertilizer, pest control and irrigation, to wheat fields infected with WSMV. This question is unanswered, largely, due to lack of information on profitability thresholds for varying levels of disease incidence and severity (Almas et al. 2016). The farmer's profit maximization problem, taking into account the level of WSM infection, can be represented by the equation

$$\max_{\theta \in \mathbb{R}^+} E\pi = \left[P^w E(y) * (1 - D) + (P^b * BM_{(y)}) * (D) - TVC_{(D)}\right] \tag{2.1}$$

$$\text{s.t}$$
$$\mathrm{E}(y) = \beta_0 + \beta_1 S_{it}$$
$$D = I(S \geq \theta); TVC_{(D)} = f(D)$$

where $E\pi$ is expected profit, P^w is the average seasonal wheat price (assumed constant), θ is the WSM infection threshold, E(y) is expected wheat grain yield, which is dependent on WSM infection severity of the i^{th} plot at a particular remote sensing time t, represented by S_{it}, P^b is the price of wheat biomass (pasture), and *BM* is the amount of biomass available to graze-out, and is a function of expected grain yield. *D* is the decision whether to abandon and graze-out an infected section, which takes a value of 1 if a farmer abandons and 0 otherwise, *TVC* denotes total variable

[1] The third strategy would require knowledge of yield from an abandoned crop, which was not captured in our data. However, we include the third strategy only for conceptual completeness. Our empirical analysis focuses only on the first and second strategies.

input costs such as fertilizer, pest control, irrigation, and labor for producing wheat ($/acre), which varies depending on *D*, while β_0 and $\beta_1$are parameters to be estimated.

One important assumption regarding the decision to abandon is that a farmer would only cease input applications if expected return from harvesting is lower than additional input costs and the grazing value. Thus in the above formulation, we assume a farmer would graze-out the wheat if a section is abandoned, to salvage some value from the wheat biomass and offset the cost of inputs applied before ceasing additional input applications. Based on estimated timing of input applications, a farmer can save costs on fungicide, supplemental nitrogen application, some level of irrigation, and harvest, since these activities would not be necessary if, after remote sensing, the farmer decides to discontinue further input applications and graze-out the wheat.

In order to link disease severity and wheat profitability, we evaluate how a given level of WSM severity at a particular point in time impacts yield potential and use these estimates to calculate profits. We can then estimate a WSM severity threshold for that time point that would render continuation of input application unprofitable.

To estimate the threshold disease severity level, we need to calculate, for each time point, the expected disease severity at which the difference in expected profits between the abandoned and non-abandoned fields or sections is zero. This can be calculated from the following net revenue functions:

(2.2) $$\mathrm{E}(\pi_1) = P^w * \mathrm{E}(y_i) - TVC_1$$

(2.3) $$\mathrm{E}(\pi_2) = VBM_t - TVC_2,$$

$$BM_t = \mathrm{E}(y_i) * 2.5$$

$$VBM_t = BM_t * P^b$$

where $\mathrm{E}(\pi)$is expected net revenue, P^w, P^b, and *y* are as defined before, TVC_1 is the variable cost of all inputs applied including grain harvest. BM_t is the quantity, in pounds, of wheat forage herein referred to as biomass, at time of remoting sensing, *t*, the value 2.5 is a conversion factor from grain yield to biomass as suggested by Xue et al. (2006). VBM_t is value of biomass, P^b denotes price of biomass, which is represented by the value of per pound weight gain of cattle grazed on wheat biomass. TVC_2 is total variable cost of inputs applied only up to the time of remote sensing. Equating (2.2) and (2.3) we obtain;

(2.4) $$P^w * E(y_i) - TVC_1 = VBM_t - TVC_2$$

Data

Data for this study were obtained from a field experiment conducted in the 2015-2016 wheat season in Dalhart area, Dallam County, Texas. To control for location and time effect, a second experiment was set up in Bushland County, Texas in the 2013-2014 season. Unfortunately, the second experiment was hailed out and so no yield was recorded. Thus, only data from the first experiment was used in this study. The experiment was conducted on a 118 acre field, which was planted to the cultivar TAM 304 on November 6th, 2015, and was under center-pivot irrigation. The following inputs were applied, fertilizer (urea) 150 lb N8 -13 -12, 30.16 inches of irrigation, herbicide (2-4 D) at a rate of 1 pint per acre, pesticide (Chlorpyrifos) also at 1 pint per acre, and fungicide (Prothioconazole and Tebuconazole). WSMV severity assessment was conducted in this field by first establishing a transect, running from one side of the field to the other. The field contained a total of 113 sampling plots (measuring $1m^2$), established across a transect, with sampling intervals ranging approximately from 4 – 10 m. However, only 99 plots had usable data, as the first 14 plots, which were located along the edge of the field, already had extremely high infection severity at the time of the first remote sensing (April 27)[2]. The length of the transect and sampling intervals were determined based on disease severity gradient from the edges of the field.

[2] Excluding these data from the analysis resulted in a lower mean reflectance reading for April, and thus increased the responsiveness of grain yield to WSM infection level for April readings. This is because for a

When wheat reached growth stage 5 – 6 on the Feekes scale (Large 1954), severity of WSM in each $1m^2$ area (5 rows) was measured by taking reflectance readings (scanning) with a hand-held hyperspectral radiometer. Previous studies (Workneh, et al. 2009) demonstrated high correlation between severity of WSM and leaf reflectance at 555 nm, so this reflectance wavelength was used as a quantitative measure of disease severity. Remote sensing of WSM was done on three dates; April 27, May 4, and May 10. Symptomatic leaves from 62 randomly selected plots were collected and tested for WSMV, *Triticum mosaic* virus (TriMV), *High Plains wheat mosaic* virus (HPWMoV), and *Barley yellow dwarf* virus (BYDV). This was done to ensure that the observed symptoms were due to WSMV since the wheat curl mite also transmits other viruses. All the 62 symptomatic samples tested positive for WSM, with only 6.5% of the samples testing positive for TriMV (in association with WSM). None of the samples tested positive for HPWMoV or BYDV, indicating WSM was by far the main disease in the field and the cause of observed symptoms.

Information on input costs and timing of applications were obtained from the farmer where the experiment was conducted, and supplemented with Oklahoma State University Department of Agricultural Economics Extension wheat budgets data. At the time of the first remote sensing, all inputs except fungicide and irrigation were fully applied. Fungicide was applied in mid-May, and irrigation continued until early June. In each of the 5-row plots, the three center rows ($0.6\ m^2$) were harvested for yield. Grain from each plot was hand-harvested on June 29, 2016, threshed and weighed to determine yield per plot. Grain yield per plot was used to determine yield in bushels per acre and Table 2.1 presents a summary of descriptive statistics of the reflectance values at different remote sensing dates and corresponding final yield.

---

given yield level, the mean April reading for the 99 plots was lower than the reading with all the 114 data points. Overall, exclusion of these data points affected only the April estimates.

Procedure

We used regression analysis to estimate the effect of WSM infection severity on yield. The general formulation of the model is

(2.5) $$lny_i = \beta_{0t} + \beta_{1t}S_{it} + \nu_{it},$$

where y_i is the wheat yield (bu/acre) in the i^{th} plot, S_{it} is reflectance value measured from each plot at time *t* (*t*= 1 for April 27th , *t*=2 for May 4th, and *t*=3 for May 10th), β_{0t} and β_{1t} are coefficients, and $\nu_{it} \sim N(0, \sigma_t^2)$ is the stochastic error term.

Accurate prediction of final yield based on reflectance reading is critical to determining disease severity threshold. Model selection and misspecification tests were conducted to help select a model with the best fit. To evaluate the correlation between yield and reflectance values, a scatter plot of yield against each of the reflectance values was graphed (Figure 2.1). The fitted line and goodness of fit test for all three reflectance readings suggest a log-linear relationship[3], implying an exponential decline in yield as WSM severity increased, consistent with other studies (Workneh et al. 2009; Byamukama et al. 2014; Almas et al. 2016).

Empirical Model and Estimation

Based on the graphed yield and reflectance values, a log-linear regression model was estimated for all three reflectance reading dates, using the specification shown below;

(2.6) $$\ln y_i = \beta_{0t} + \beta_{1t}S_{it} + \nu_{it}$$

where all variables and parameters are as defined before. Since the natural logarithm is a nonlinear transformation, expected yield cannot be calculated by simply setting the error term to zero. Since yield follows a lognormal distribution, expected yield is

(2.7) $$\widehat{y_{it}} = e^{\widehat{\beta_{0t}} + \widehat{\beta_{1t}} S_{it} + \left(\widehat{\sigma}_t^2/2\right)}$$

[3] As a robustness check, we fitted alternative functional forms, including linear, double-log, and linear-log. Results of these estimates are available upon request. Generally, the log-linear forms provided a better fit.

where $\hat{\sigma}_t^2$ is the sample variance of the error term, and all other variables and parameters are as defined before, with the hat on parameters indicating that the parameters are estimated. Substituting (2.7) in equation (2.4) and solving for the threshold, *S*, at time *t*, we obtain;

$$S_{it} = \frac{\ln\left[\frac{VBM - TVC_2 + TVC_1}{P_W} - \left(\hat{\sigma}_t^2 / 2\right)\right] - \hat{\beta}_{0t}}{\hat{\beta}_{1t}} \tag{2.8}$$

where all variables and coefficients are as defined before.

Our analysis proceeded in two stages. The first stage involved regression analysis to estimate yield response function to WSM, and the second stage used the coefficients from the estimated yield function in the partial budget analysis to calculate the threshold values. We used nonparametric bootstrapping, with 1000 replications, to obtain the sampling distributions of the threshold estimates (mean, standard deviation, and confidence interval).

Misspecification Tests

We conducted model misspecification tests because misspecification can lead to biased and inconsistent estimators, resulting in inappropriate inferences (McGuirk et al. 1993). A scatter plot of reflectance readings, against wheat yield suggested an exponential yield-reflectance reading relationship (Figure 2.1), thus a log-linear model was fitted. Following (D'Agostino, Belanger and D'Agostino 1990), the K^2 omnibus test of normality was conducted, and the test did not detect deviations from normality due to either skewness or kurtosis for the May 4th and May 10th sensor readings. Other tests conducted Lagrange multiplier test for heteroskedasticity (Breusch and Pagan 1980); and a joint conditional mean and conditional variance tests, using the comprehensive specification tests as suggested (McGuirk et al. 1993). None of the tests detected statistically significant misspecification.

Results and Discussion

Table 2.2 present regression estimates of effect of WSMV infection severity (reflectance readings) on wheat yield, for all three reflectance reading dates. All three fitted regressions predicted final yield relatively well, as indicated by relatively narrow confidence intervals, small standard errors, and high R-squared values. Of the three reflectance data collection dates, the May 10 readings predicted yields more accurately, with an R-squared value of 0.78, followed by the April 27 readings at 0.71, and lastly the May 4 readings with R-squared of 0.70. The coefficients on reflectance readings – representing WSM infection severity - are negative and statistically significant in all three regression models, confirming the negative relationship between WSM severity and final yield. In terms of magnitude, holding all else equal, a one unit increase in the April 27 reflectance value corresponded to a 30 percent reduction in expected final yield, while a similar increase in the second and third reflectance values corresponded to yield reductions of 31 and 35 percent, respectively.

The difference in magnitude across time (i.e., across the three sets of reflectance values) can be attributed to increased infection severity over-time, since the same plots had reflectance readings taken at three different time points. Therefore, at lower levels of infection severity (early in the season), an increase in infection severity has lower effect on expected final yield, compared to a similar increase later in the season, when infection severity is high, such that a marginal increase in infection would result in relatively higher yield loss. This finding is consistent with previous reports (Workneh et al. 2017; Almas et al. 2016; Byamukama et al. 2012). However, this may not be the case if the same level of infection was observed at different times. For example, Pradhan et al. (2014) inoculated wheat with the same level of WSM at different developmental stages and found higher yield impacts for earlier infection than later.

After estimating the WSM – yield potential relation (yield predictor), we conducted a partial budget analysis to evaluate economic threshold yield potential. We then estimated the economic WSM infection threshold as the level of infection corresponding to the threshold yield potential for each remote sensing date. We then compare returns from two production decisions namely: 1) continue input applications and harvest grain at the end of the season, and 2) cease input applications mid-season and graze-out. Estimates of graze-out revenue were obtained by converting expected grain yield to biomass using a conversion factor of 2.5, suggested by Xue et al. (2006). As mentioned earlier, the value of biomass is represented by the value of beef cattle gain (value of gain)-determined as the difference in the weight of beef cattle before and after placement on wheat pasture.

Our value of gain estimate draws from previous analyses, such as Belasco et al. (2009) and Tumusiime et al. (2011), who estimated dry matter feed conversion into beef gain. Tumusiime et al. (2011) assume that for every 10 pounds of wheat forage biomass consumed, a steer gains 1 pound, and each pound of gain is valued at $0.45. Their analysis was based on WSM-free wheat forage during winter months. However, forage quality from WSM infected crop is expected to be low and the crop is approaching senescence, and thus, cattle would have to consume relatively higher quantities of the "*poor-quality*" forage per pound of gain. Further, during April and May (period of interest for our analysis), the price of wheat forage would adjust downwards to reflect abundance of grass forage at that time of the year. To account for quality decline, we adjusted the pounds of forage required for a pound of gain from 10 to 20. The value of gain, reflects an adjusted price of forage, downwards from $0.45/lb to $0.35/lb to reflect availability of substitute forage during that time of the year as well as potential costs of moving cattle or electric fencing to fence off part of a field. Albeit arbitrary, to some extent, these adjustments partially help account for changes in forage quality, as well as capturing seasonal fluctuations in forage price. Using these assumptions we calculated revenue from graze-out as (BM/20)*0.35, which gives a per pound biomass price of $0.035.

One important estimate from partial budget analysis is the total variable costs a producer could potentially save in inputs by the time of remote sensing, which is the same time that a decision whether to cease input application and graze out is to be made. Based on information on timing of input application, at the time of remote sensing (April 27th –May 10th), producers could save $60.29/acre in inputs and irrigation, if they decide to abandon the crop and graze it out due to low yield potential. This amount also represents cost savings as a result of the decision to stop input application and graze-out the biomass mid-season, when potential grain yield is low. Using the above-mentioned assumptions, we estimated the economic threshold yield by setting net returns from grain harvest equal to returns from graze-out. This gave us a threshold grain yield of 29.7 bu/acre. At this yield level, net returns from grain harvest and graze-out are equal (Figure 2.2).

The WSM infection threshold was then calculated as the level of severity that corresponds to a grain yield potential of 29.7 bu/acre. Following estimation of economic threshold yield, we used the estimated yield predictor and partial budget estimates to calculate the economic infection severity threshold for each remote sensing date. Specifically, we used equation (1.8) to estimate threshold values for all three remote sensing dates, from 1000 bootstrapped samples. Table 2.3 presents a summary of threshold estimates and their sampling distributions. It should be noted that the threshold estimates presented in Table 2.3 are a function of input and output prices. Therefore, the threshold will vary not only depending on remote sensing dates, but also contingent on input and output prices used to construct the partial budget. Results indicate varying WSM severity threshold levels for all the three remote sensing dates. The May 4 readings had the highest threshold value at about 9.5, followed by May 10 with a threshold value of 8.4, and in the third place was the April 27 readings at 8.3. This result is expected, except for May 4 having a higher threshold than the May 10 readings.

As the severity thresholds correspond to a point where net returns from graze-out and grain sales are equal (Figure 2.2), the estimated thresholds should be regarded as a signal to consider graze-out. Our threshold analysis indicated potential to save resources by discontinuing mid-season input applications and introducing cattle for grazing, in about 14% of the sampled plots. The difference in thresholds values between the May 4 and May 10 is inconsistent with most studies on crop disease infection and final yield (e.g., Hunger et al. 1992; Price 2015) that suggest lower threshold for earlier infection than later. A plausible explanation for this finding is that high disease severity early in the season (May 4 in our case) could have resulted in senescent tissue, that was not yellow but brown, thus a lower reading at 555nm for the latest (May 10) readings. It could also be the case that there was background contamination from bare soil due to loss of canopy as infection severity increased further by the time of the May 10 remote sensing. Grazing may not always be an option if cattle are not nearby or the affected area is not fenced or is impractical to fence.

We conducted further analysis assuming graze-out is not possible, and a grower has to completely abandon the crop. Results of this analysis indicated the threshold yield to be 12.9 bu/acre. The three reflectance reading thresholds for this scenario are 11.035 for April 27 remote sensing, about 12 for May 4, and 10.6 for May 10 remote sensing (Table 2.5). At these thresholds, only 1 percent of the sampled plots could save resources by abandoning the infected crop.

Detecting the disease early enough in the season is critical for making mid-season management decisions. The farmer would have had more opportunity to adjust inputs if the disease had been detected earlier than it was in this study. Ideally, farmers should determine disease incidence and severity between late-March to mid-April, and if reflectance readings (disease severity) exceed the threshold, then farmers should consider ceasing input applications and graze-out the wheat biomass. However, the best strategy is for farmers to prioritize good management practices that eliminate the green bridges. Although wind may carry mites from one field to another, elimination of the green bridges may prevent and/or reduce chance of early infection.

Conclusion

This study used field experiment data from a farmer's field, to estimate WSM severity threshold beyond which continued application of inputs is uneconomical. Results of the analysis indicated exponential yield decline with increasing WSM infection severity. For example, our estimates show that by April 27, an increase in reflectance readings by one corresponds with a yield decline of 30 percent, while a similar increase by May 4 and 10 corresponds with 31 and 35 percent yield decline, respectively. This result indicates how rapidly WSM severity progresses, and how much yield can potentially be lost for a given level of severity during the season. Given the relatively quick disease infection progression, for optimum results, farmers should aim to conduct remote sensing to quantify disease severity, as early in the season as possible after the crop comes out of dormancy. This suggestion is supported by data collected in this study. When reflectance measurements were first collected April 27, the first 14 plots along the edge of the field were already severely infected and had an average reflectance reading near 12. These plots were in an area of the field that represented the initial disease introduction area; an area that was likely first infected early in the fall, soon after the crop emerged. It is highly probably that disease symptoms were visible in the fall and certainly could have been observed easily by late March. However, by the April reading, disease severity was so severe that subsequent reflectance readings at 555nm (yellow region of the spectrum) no longer adequately represented disease severity, because infected tissue was senescing and turning brown.

Our results are consistent with those of others who have attempted to estimate the effect of WSM on yields. However, our study represents the first analysis to estimate the economic threshold of the disease, during the growing season. Our estimates indicated the threshold reflectance to range from about 8.3 to 9.5, for readings taken around late April to early May, and are sensitive to input and output prices used in the construction of partial budgets. With well over a tenth of the sampled plots having reflectance values greater than the threshold, results suggest farmers may potentially save resources and salvage some value by discontinuing input applications and grazing out the wheat forage from infected fields whose infection severity exceeds the estimated threshold values at a particular time. However, farmers need to continue to prioritize good management practices such as clearing the field of volunteer wheat and weeds early enough before planting, to eliminate the green bridge, and reduce chances of infection.

In this study, we have attempted to quantify WSM effects on yield, and severity threshold using data from one field experiment for a single year, and were thus unable to control for year, cultivar and location effects. More precise estimates could be obtained from future studies that obtain data from multiple years, cultivars, and locations. In addition, future studies should consider obtaining data on yield, and input costs from grazed-out fields as a counterfactual to grain-harvested fields, with and without additional inputs. Furthermore, given that mite vectored virus disease symptoms typically begin to show up in late March, it is critical that future studies begin sensing early to avoid plant tissue senescence which negatively impacts the correlation between reflectance values at 555 nm and disease severity.

References

Almas, L.K., J.A. Price, F. Workneh, and C.M. Rush. 2016. "Quantifying Economic Losses Associated with Levels of Wheat Streak Mosaic Incidence and Severity in the Texas High Plains." *Crop Protection* 88:155–160. Available at: http://dx.doi.org/10.1016/j.cropro.2016.06.012.

Belasco, E.J., M.R. Taylor, B.K. Goodwin, and T.C. Schroeder. 2009. "Probabilistic Models of Yield, Price, and Revenue Risks for Fed Cattle Production." *Journal of Agricultural and Applied Economics* 41(01):91–105.

Breusch, T.S., and A.R. Pagan. 1980. "The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics." *The Review of Economic Studies* 47(1):239–253. Available at: https://academic.oup.com/restud/article-lookup/doi/10.2307/2297111.

Burrows, M., G. Franc, C. Rush, T. Blunt, D. Ito, K. Kinzer, J. Olson, J. O'Mara, J. Price, C. Tande, A. Ziems, and J. Stack. 2009. "Occurrence of Viruses in Wheat in the Great Plains Region, 2008." *Plant Health Progress* July.

Byamukama, E., S. Tatineni, G.L. Hein, R.A. Graybosch, P.S. Baenziger, R. French, and S.N. Wegulo. 2012. "Effects of Single and Double Infections of Winter Wheat by *Triticum mosaic virus* and *Wheat streak mosaic virus* on Yield Determinants." *Plant Disease* 96(6):859–864.

Byamukama, E., S.N. Wegulo, S. Tatineni, G.L. Hein, R.A. Graybosch, P.S. Baenziger, and R. French. 2014. "Quantification of Yield Loss Caused by Triticum Mosaic Virus and Wheat Streak Mosaic Virus in Winter Wheat Under Field Conditions." *Plant Disease* 98(1):127–133.

D'agostino, R.B., A. Belanger, and R.B. D'Agostino Jr. 1990. "A Suggestion for Using Powerful and Informative Tests of Normality." *The American Statistician* 44(4):316–321.

Department of Agricultural Economics, Oklahoma State University. 2017. “Irrigated Wheat Enterprise Budget - Grain and Graze.” Available at: http://agecon.okstate.edu/budgets/sample files/Wheat2.1 irr ph 2.pdf

Griffin, T., and J.M. Lowenberg-DeBoer. 2017. "Impact of Automated Guidance for Mechanical Control of Herbicide-Resistant Weeds in Corn." *Journal of Applied Farm Economics* 1(2):62-74.

Kuosmanen, T. 2006. “Specification and Estimation of Production Functions Involving Damage Control Inputs : A Two-Stage, Semiparametric Approach.” *American Journal of Agricultural Economics* 88(2):499–511.

Large, E. 1954. Growth Stages of Cereals: Illustrations of the Feek’s scale. *Plant Pathol*. 3:128-129

Mbah, M.L.N., G.A. Forster, J.H. Wesseler, and C.A. Gilligan. 2010. “Economically Optimal Timing for Crop Disease Control under Uncertainty: An Options Approach.” *Journal of The Royal Society Interface* 7(51):1421–1428.

Mcguirk, A.M., P. Driscoll, J. Alwang, A.M. Mcguirk, P. Driscoll, and J. Alwang. 1993. “Misspecification Testing : A Comprehensive Approach.” *American Journal of Agricultural Economics* 75(4):1044–1055.

McMullen, M., and D. Waldstein. 2010. “Wheat Streak Mosaic.” *Plant Disease Management NDSU Extension Services* 646(September):602–604.

Michael, C.L., and W.G. William. 1993. “Survival of Wheat Streak Mosaic Virus in Grass Hosts in Kansas from Wheat Harvest to Fall Wheat Emergence.” *Plant Disease* 77(3):239–242.

Pradhan, G., Xue, Q., Jessup, K., Hao, B., Price, J., and Rush, C. M. 2015. Physiological Responses of Hard Red Winter Wheat to Infection by *Wheat Streak Mosaic Virus*. *Phytopathology* 105: (5) 621-627.

Price, J. Ecology and Epidemiology of *Wheat Streak Mosaic Virus, Triticum mosaic virus , and their Mite Vector in Wheat and Grassland Fields*. Texas Tech University. Available at:

https://ttu-ir.tdl.org/ttu-ir/bitstream/handle/2346/62348/PRICE-DISSERTATION-2015.pdf?sequence=1.

Tumusiime, E., B.W. Brorsen, J. Mosali, J. Johnson, J. Locke, and T.J. Biermarcher. 2011. “Determining Optimal Levels of Nitrogen Fertilizer Using Random Parameter Models Effects of Government Policy on Agricultural Markets View project Forage Production and Nutritive Value View project Emmanuel Tumusiime.” *Agricultural and Applied Economics* 43(4):541–552.

Velandia, M., R.M. Rejesus, D.C. Jones, J.A. Price, F. Workneh, and C.M. Rush. 2010. “Economic impact of Wheat Streak Mosaic Virus in the Texas High Plains.” *Crop Protection* 29(7):699–703.

Workneh, F., D.C. Jones, and C.M. Rush. 2009. “Quantifying Wheat Yield Across the Field as a Function of Wheat Streak Mosaic Intensity: A State Space Approach.” *Phytopathology* 99(4):432–440.

Workneh, F., S. O’Shaughnessy, S. Evett, and C.M. Rush. 2017. “Relationships Between Early Wheat Streak Mosaic Severity Levels and Grain Yield: Implications for Management Decisions.” *Plant Disease* 101(9):1621–1626.

Xue, Q., Z. Zhu, J.T. Musick, B.A. Stewart, and D.A. Dusek. 2006. “Physiological Mechanisms Contributing to the Increased Water-use Efficiency in Winter Wheat under Deficit Irrigation.” *Journal of Plant Physiology* 163(2):154–164.

Table 2.1. Descriptive Statistics for Three Reflectance Values and Assessment Dates and Yield

Variable	Mean	Std. Dev.	Minimum	Maximum
First reflectance (April 27)	6.11	1.73	3.83	11.63
Second reflectance (May 4)	7.19	1.66	4.78	11.96
Third reflectance (May 10)	6.74	1.57	4.13	11.45
Yield (bu/acre)	65.22	27.08	6.79	109.27

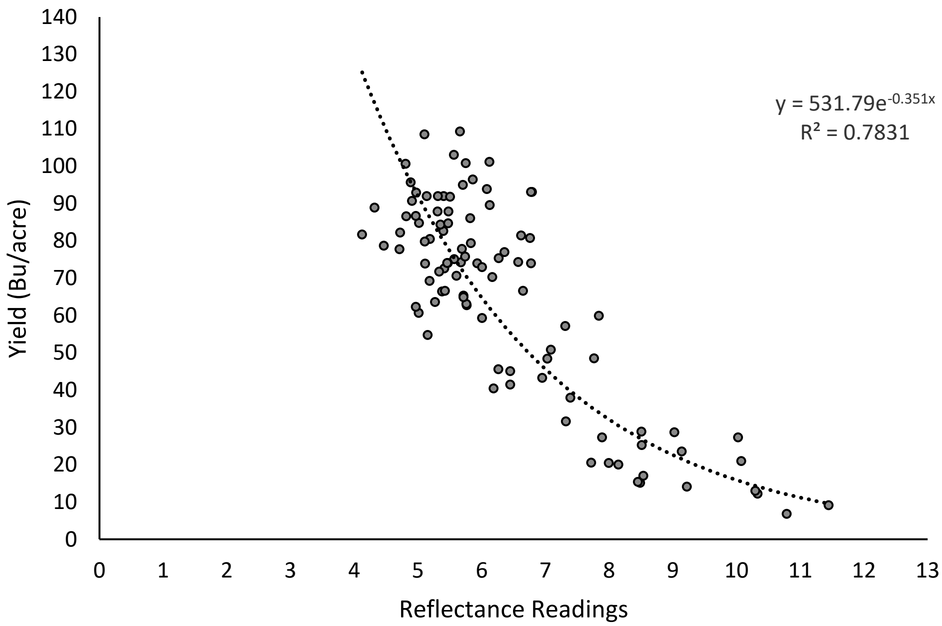


Figure 2.1. Scatter Plot of the May10th Reflectance Reading and Wheat Grain Yield (bu/acre)[4]

[4] Scatter plots for the other two reflectance readings versus yield, showing a similar relationship, were generated, although not presented here.

Table 2.2. Regression Estimates of the Effect of WSMV on Wheat Yield (log bu/acre) at the April 27, May 4, and May 10 Remote Sensing

Estimate	27-Apr	4-May	10-May
Intercept	10.009**	10.502**	10.487**
	(0.125)	(0.154)	(0.124)
Coefficient	-0.303**	-0.313**	-0.351**
	(0.020)	(0.021)	(0.019)

Note: Standard errors in parenthesis; *$p < 0.05$, and **$p < 0.01$. Number of observations = 99.

Table 2.3. Partial Budget - Decision to Graze-out from a WSMV Infected Crop

Item	$/Acre	Item	$/Acre
Added income due to graze-out:		**Added costs due to graze-out:**	
Graze-out revenue (224 lbs of gain at $0.35/lb of gain)	78.05	Cost of grazing	0.00
Reduced costs due to graze-out:		**Reduced income due to graze-out:**	
Fungicide	19.00	Wheat grain yield (29 bu/acre) at $4.67/bu	138.69
Irrigation (25% of total irrigation cost)	18.75		
Harvest (Machine + Labor)	22.54		
Subtotal	138.69	Subtotal	138.69
Net change: 138.69 -138.69 =	**0.00**		

Table 2.4. WSMV Infection Threshold by Remote Sensing Date under Graze-out Scenario

Remote Sensing Date	Threshold Estimate	SE	95% Confidence Interval	
			Lower Limit	Upper Limit
April 27th	8.311	0.264	8.262	8.360
May 4th	9.522	0.296	9.467	9.577
May 10th	8.390	0.172	8.358	8.422

Note: Results based on 1000 nonparametric bootstrapped samples from 99 sampled plots.

Table 2.5. WSMV Infection Threshold by Remote Sensing Date under Complete Abandonment Scenario

Remote Sensing Date	Threshold Estimate	SE	95% Confidence Interval	
			Lower Limit	Upper Limit
April 27th	11.035	0.490	10.944	11.126
May 4th	11.988	0.523	11.891	12.085
May 10th	10.634	0.302	10.578	10.690

Note: Results based on 1000 nonparametric bootstrapped samples from 99 sampled plots.

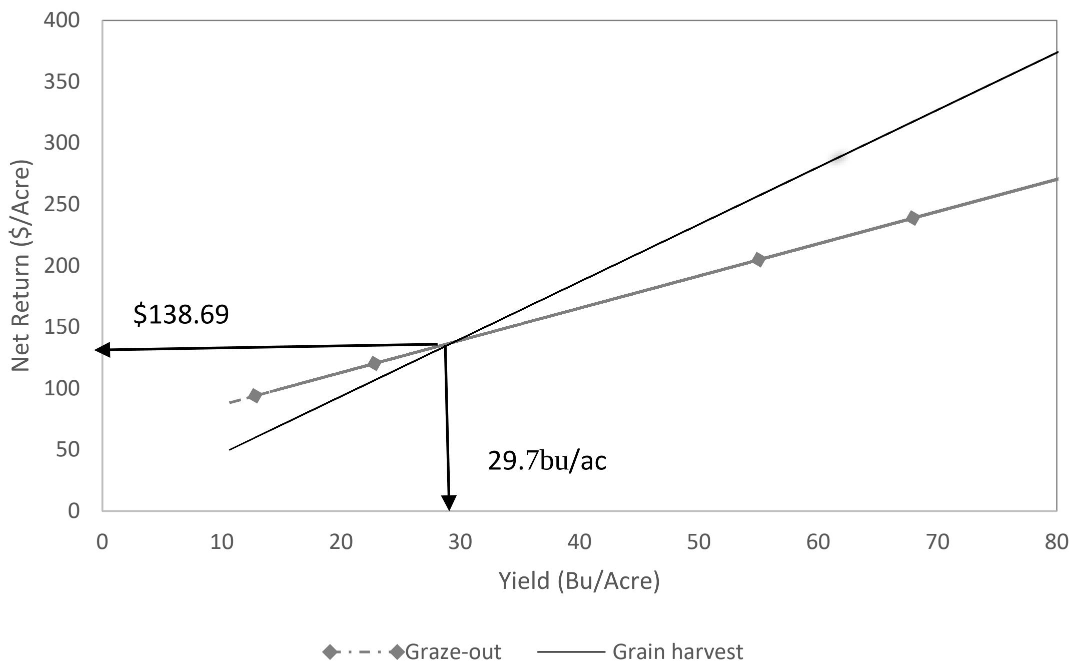


Figure 2.2. Economic Threshold Yield and Returns for Wheat Streak Mosaic Infected Crop

CHAPTER III

IMPROVING COST OF GAIN PREDICTION FOR FEEDER CATTLE

Introduction

Feedlot operator decisions regarding the size of feeder cattle to purchase and place in the feedlot depend, in part, on the expected cost of gain (COG). Cost of gain refers to the average cost per pound of weight an animal gains after it is placed on feed. Feedlots use cost of gain to evaluate the tradeoff between purchasing lighter or heavier feeder cattle. Generally, this tradeoff is between two decisions; 1) purchase more pounds via heavier feeders and spend less on feed to bring the cattle to slaughter weight; or 2) purchase lighter feeders and put more weight on in the feedlot, which implies higher feeding costs to bring the animal to slaughter weight. Hence, cost of gain can be used to estimate a feedlot breakeven price[5] for lighter weight feeders compared to heavy feeders.

Monthly data on ex-post COG estimates for current feedlot closeouts (i.e., the feeding period just finished) is useful to explain the relationship between feeder prices across weights and estimated feedlot purchase breakeven price. In fact, it is often used as a naïve estimate for future COG. However, it may be possible to improve understanding of this relationship by constructing a better ex-ante COG prediction model. There is, for example, seasonal variation in the relationship between feeder prices and feedlot breakeven price by weight.

[5] Feedlot breakeven price in this study refers to feedlot purchase breakeven price, and not the bid price.

An improved COG prediction model could facilitate more accurate estimates of expected COG, and thus corresponding feedlot purchase breakeven price, for lighter animals relative to heavier animals.

The copious literature on feeder cattle price and COG explains the relationship between corn price, COG, and the resultant effect on feeder market and breakeven prices. Albright, Schroeder and Langemeier (1994) analyzed the COG and corn price relationship in two Kansas feedlots and estimated that, on average, 64 percent of variability in cost of gain could be attributed to corn price variability. Thus, corn price has a strong relationship with COG and feedlot profitability, as corn typically comprises the largest proportion of most feedlot rations (Mark, Schroeder and Jones 2000; Anderson and Trapp 2000b; MacDonald and Schroeder 2003). Anderson and Trapp (2000a) note that the strong relationship between corn price and COG led most popular publications to assume unit elasticity for COG with respect to corn price (e.g., Fox 1996; Maday 1996), implying that COG would change by the same percentage as corn price. This assumption may lead to inaccurate COG projections, given the dynamic nature of the cattle feeding process. For example, an increase in corn price may result in substitution of other feed stuffs for corn (e.g., wheat), and changing feed compositions. Further, Langemeier, Schroeder and Mintert (1992) and Marsh (1999) note that a change in corn price will prompt feedlot operators to change the weight of cattle placed on feed, hence affecting COG. These dynamics affect how COG responds to corn price variation and should be accounted for in COG projections.

The ex-post COG for the most recent feeding period does provide useful information for estimating feedlot breakeven price for cattle placed in the current month. However, the use of ex-post COG as expected COG could preclude the detection of seasonal variation in the relationship between corn prices, COG, and feeder prices for different feeder weights. Existing studies (Maday 1996; Anderson and Trapp 2000b; Anderson and Trapp 2000a; MacDonald and Schoeder 2003) provide important insight into the relationship between corn price, COG, feeder prices, and feedlot profitability. However, most publications (Kansas State University Agricultural Experiment Station 2016; Albright, Schroeder, and Langemeier 1994) use a simple regression of profit or COG on corn futures price and feeder cattle price, among other variables, to predict cost of gain. Such a model implicitly assumes a static relationship between COG and corn prices – which may not always hold. One notable exception is Anderson and Trapp (2000a), who model the cost of gain-corn price relationship as an autoregressive process using monthly time series data aggregated at the national level spanning 1980 through 1996. After controlling for seasonality they determine that the COG response with respect to corn price change is less than the commonly assumed unity elasticity and that breakeven estimates based on COG assuming a unity elasticity are inaccurate. However, note that their use of nationally aggregated data will not capture regional variations in terms of feeding dynamics, availability of corn substitutes, and market conditions.

The current analysis builds on previous studies such as Anderson and Trapp (2000a) using regional data, to construct a COG prediction model that incorporates the dynamics of the COG - corn price relationship and seasonality to more accurately predict expected COG and thus feedlot breakeven purchase prices. Since feedlot operators mostly rely on projected COG estimates when making feeder cattle purchasing decisions, more accurate COG predictions could improve their breakeven price estimates and facilitate better evaluation of the tradeoff between lighter and heavier feeders. Findings of this study are also expected to help evaluate to what extent the difference between the current ex-post COG (naïve model) and predicted COG estimates explain the observed seasonal variations in the relationship between feeder market prices and feedlot breakeven price by weight.

Against this background, the objectives of this study are twofold: 1) to predict feeder cattle cost of gain; and 2) estimate feedlot breakeven price across weight groups based on predicted cost of gain, and evaluate its spread with respect to feeder market price.

Cattle Feeding Profitability and Feedlot Breakeven

Research on cattle feeding and profitability (Zhao, Du and Hennessy 2011; Belasco et al. 2009; Dhuyvetter and Schroeder 2000; Ash 1994) has shown that COG comprises one of the most important components of feedlot profitability. All costs incurred by feedlots in the process of adding weight to feeder cattle, such as feed costs, veterinary expenses, yardage fees, interest charges, and death loss are categorized as cost of gain (Albright, Schroeder and Langemeier 1994). However, feed costs represent the most significant percentage, and typically the most variable cost component, making it an important influence on cattle feeding profitability (Marsh 1999; Albright, Schroeder and Langemeier 1994). Besides feed, another major input in cattle feeding is of course feeder cattle, which feedlots purchase mostly from stockers/backgrounders. Once purchased, feeder cattle are placed on a high-energy grain diet, comprised mostly of corn.

During this feeding period, which typically lasts about 150-210 days (Zhao et al. 2011; Anderson and Trapp 2000a), cattle are fed to reach the desired slaughter weight before being sold to beef packers. Depending on market conditions, such as corn and feeder cattle prices, a feedlot operator may purchase heavier feeders or purchase lighter feeders and put more weight on in the feedlot. Generally, price per pound of lighter-weight feeder cattle is higher than for heavier feeders, because feeding cost per pound of gain for lighter feeders is lower than the selling price per pound.

A number of empirical studies (e.g., Belasco, Ghosh and Goodwin 2009; Dhuyvetter and Schroeder 2000; Albright et al. 1994) have shown that healthy lighter-weight feeders tend to have higher feeding efficiency than heavier feeders; hence lighter feeders require less corn per pound of gain than heavier feeders. However, lighter feeders require more days on feed and total feed quantity to reach slaughter weight relative to heavier feeders (Dhuyvetter and Schroeder 2000). Belasco et al. (2009) and Belasco, Ghosh, and Goodwin (2009) point out that heavier feeders tend to have higher average daily gain and hence spend fewer days than lighter feeders on feed to reach slaughter weight. This link between feeder cattle weight and days on feed leads feedlots to increase placement weight as a cost minimization strategy when corn prices are high. This is because higher corn prices lead to higher feeding costs, and by extension, cost of gain.

Thus, although lighter feeders feed efficiently, they are worth more only if the feeding cost is low enough to offset the high cost of gain. It is evident that the decision regarding size of feeder cattle to purchase depends, in part, on cost of gain. It is also clear from the preceding discussion that feeder cattle placement weight has important implications for feedlot profitability. Unlike other decision variables such as feeder cattle price, and current corn price which are observable at time of feeder cattle purchase actual COG is not observable ex-ante. Therefore, cattle feeders must form expectations of COG based on known information about other variables such as corn price, placement weight, and season.

Theoretical Model

Because this study focuses on predicting COG ex-ante, the theoretical model presented below only includes variables observable at placement. Consider a price-taking feedlot operator, with a generalized primal profit equation

(3.1) *Profit = Revenue – Cost of Feeder Cattle – Cost of Gain –Fixed Costs*

where both revenue and costs are in dollars per hundred weight ($/cwt). If we assume that expected closeout COG over a given feeding period, *K*, can be predicted at time of feeder cattle purchase or placement (*t*), an operator’s decision choice regarding size of feeders to purchase can be represented by the following profit maximization equation:

(3.2) $$\max_F \mathrm{E}(\pi) = \mathrm{E}[P_W \cdot W(F, G(C)) \cdot (1 - DL) - P_F(F) \cdot F - COG(P_C, F, S)] - FC$$

s.t.

$$\mathrm{E}(W) = F + G(C)$$

$$0 < E(DL) < 1$$

$$C = f(F, S)$$

$$\mathrm{E}(COG) = \frac{E(P_C) \cdot C(F,S)}{E(G)}$$

$$\mathrm{E}(W) \geq E(F)$$

where $E(\pi)$ is the operator's expected profit at the end of the feeding period (t+K), E(P_W) is the expected price of fed cattle ($/cwt) at time t+K, E(W) is the expected weight of fed cattle sold to beef packers, E(G) is the expected weight gain, assumed to be a linear function and is dependent on quantity of corn (C). The quantity of corn in turn depends on placement weight, F, and placement season, S which accounts for variation in feeding efficiency. E(DL) is the expected death loss at the end of the feeding period, E(P_F) is the price of feeder cattle ($/cwt), E($F$) is the expected weight of feeder cattle at placement. E(COG) is the expected closeout cost of gain ($/cwt), conditional on expected corn price E(P_C), placement weight, and placement season (S). We assume zero covariance among expectations.

Previous empirical research on price expectations (e.g., Antonovitz and Green 1990; Eales et al. 1990; Gardner 1976) suggest futures price as an appropriate representation of expected price; thus we use a weighted corn futures price to represent expected corn price. Seasonal variables capture the variation in feeding efficiency. Generally, feeding efficiency decreases during cold months and increases during warm months, thus affecting COG (Belasco, Ghosh, and Goodwin 2009; Mark and Schroeder 2002). However, it should be noted that animals eat less in extreme hot weather and gain tends to be negatively affected. FC is the operator's fixed cost per given amount of weight gain and is assumed negligible relative to feed costs, typically an upper bound of 10% of total cost (Ellis et al. 2009; Zhao et al. 2011). The first order condition of profit maximization equation (3.2) with respect to placement weight *F* yields the following:

$$(3.3) \qquad \frac{\partial \mathrm{E}(\pi)}{\partial \mathrm{E}(F)} = P_W \cdot \frac{\partial W}{\partial F} - \frac{\partial W}{\partial F} \cdot P_W \cdot DL - \frac{\partial P_F}{\partial F} \cdot F + P_F - \frac{\partial \mathrm{E}(COG)}{\partial \mathrm{E}(F)}$$

Equation (3.3) above shows a profit maximization function that is responsive to expected placement weight E(*F*) and expected cost of gain E(*COG*).

Thus a profit maximizing, price-taking feedlot will maximize profit by producing at the point where marginal product value, P_W ($/cwt) equals input marginal cost, where the latter is comprised feeder cattle price (P_F) and *COG*, a function of corn price and placement weight. Because a feedlot is assumed to be a price-taker, the only choice variable in this theoretical model is placement weight -the weight of lighter feeder cattle to be purchased and placed on feed. Given that COG is a function of placement weight, a feedlot's COG will vary depending on the size of feeder cattle placed on feed. Thus varying placement weight translates in varying COG, which in turn affects profit. Since feedlots do not know actual COG until closeout, estimating expected closeout COG in advance of placement can inform feedlots' cattle purchase decisions regarding feeder placement weights.

Based on equation (3.2) and the first order condition (3.3), expected COG can be modeled as shown below:

$$(3.4) \qquad \mathrm{E}[COG] = f\,(P_C, F, S).$$

Expected COG can then be used to calculate the current feedlot breakeven purchase price for lighter weight feeders (550-600 pound, 600-650 pound, 650-700 pound, 700-750 pound, and 750-800 pound) relative to the current price of heavy feeder cattle, which we assume to be 800-850 pounds (lb) in this study. The calculated breakeven price can then be compared to the current feeder market price for lighter cattle to determine the price spread for lighter feeders. The breakeven price can be derived from the equilibrium condition expressed as *Value of Lighter Feeder + Cost of Gain (Gain) = Value of Heavy Feeder.* Mathematically, this expression is represented as

$$P_L \cdot W_L + \mathrm{E}[COG](W_H - W_L) = P_H \cdot W_H \quad (3.5)$$

where P_L is the price for a lighter feeder, P_H is price of a heavier feeder and is known, W_H is the weight of a heavy feeder animal, and W_L is the weight of a lighter feeder being evaluated against a heavy feeder. The feedlot breakeven purchase price (BE_L) at which equation (3.5) holds true is obtained by solving for P_L (and replacing P_L with BE_L) to yield the following:

$$BE_L = \frac{P_H \cdot W_H - \mathrm{E}[COG] \cdot (W_H - W_L)}{W_L} \quad (3.6)$$

Thus, BE_L represents the price at which an operator is indifferent between purchasing a lighter feeder or a heavy feeder (i.e., willingness to tradeoff lighter feeder and heavy feeder cattle).

The price spread is then estimated as the price difference between calculated breakeven price (BE_L) for lighter feeders and feeder market price for lighter feeder cattle:

$$P_{sL} = BE_L - P_L^O \quad (3.7)$$

where P_{SL} is the price spread for a given lighter weight feeder, and P_L^O is the observed current market price for a given lighter weight feeder animal.

Typically, all variables except COG in equation (3.5) are known at the time of placement. Thus it is expected that accurate COG prediction constructed in this study will improve the accuracy of breakeven price estimates obtained from equation (3.6) relative to the breakeven obtained from using current closeout COG as a projection for current placement.

Data

Data for this study is drawn from the Kansas Feedlot Performance and Feed Cost surveys, which are aggregated to provide monthly summaries. Data for the period spanning January 1992 to June 2017 are from the Livestock Marketing Information Center (LMIC). The monthly summary contains feedlot performance and closeout data from Kansas commercial cattle feeding operations. Data are collected via a monthly survey of a number of Kansas feedlots with the number of respondents varying by month. Data include monthly averages of closeout COG for all pens of cattle finished in a given month, days on feed, monthly corn futures price for all closing months, placement weight, days on feed, and average daily gain. More specifically, all the above-mentioned variables are averages over all cattle finished that month, regardless of placement date. In this study we focus on steers. Data used to calculate spreads between breakeven price and market prices are weekly averages from Oklahoma cattle auctions.

The aggregated nature of the LMIC dataset provides average cost of gain for the entire feeding period for all cattle slaughtered in a particular month. However, placement date and month may vary among cattle slaughtered in a given month, similar to the data in Anderson and Trapp (2000). Although the LMIC data do not provide placement date, they does provide average days on feed for all cattle slaughtered in a given month. Thus it is possible to estimate placement date (month) by taking the time difference between closeout date and days on feed (in months). A summary of variables of interest in the dataset is given in Table 3.1. In addition, graphs of key variables by season are given in appendices section (Figures 3.1A-3.5A).

The data provide COG estimates for the entire feeding period for cattle slaughtered in a given month. Thus, an accurate expected corn price should be representative of the entire feeding period. Because feedlots typically procure corn through forward contracting and keep inventories for future use, past corn prices most likely have an effect on COG (Anderson and Trapp 2000a). We address this by calculating a weighted expected corn price that adjusts for how frequently a corn futures price is referenced during the feeding period based on specific months on feed and the associated corn futures closeout months. For example, if cattle are placed in January and finished in May, then three corn futures contracts prices quoted at time of placement in January will be referenced during this feeding period, i.e. March, May, and July. The March contract will be referenced twice (January and February); the May contract will also be referenced twice

(March and April); and the July contract will be referenced once (May). In this case, March and May corn futures will each be allocated a weight of 2/5, whereas the July futures will be allocated a 1/5 weight. Therefore, the corn futures price weight can be calculated as:

$$(3.8) \qquad H_m = \frac{R_m}{K}$$

where H_m is the weight for a given futures contract price, R_m is the number of times a particular corn futures contract price for month *m* (*m* = March, May, July, September, December) is referenced during the feeding period (*K*) assumed to be constant (5 months). Using this weight, we can estimate expected corn price, $E[P_C]$, for a given feeding period as a sum of the weighted corn futures price relevant to the feeding period as quoted at time *t*. Mathematically, this is expressed as

$$(3.9) \qquad E[P_C] = \sum_{m=1}^{5} P_m^f * H_m$$

where P_m^f is the corn futures contract price for a given contract expiration month.

Procedure and Estimation

Stationarity and Cointegration Tests

In time series data, non-stationary of variables is frequently observed (Wooldridge 2008; Greene 2002). Stationarity in time series analysis is desired as it assures stability of the relationship between variables. If either corn price or cost of gain or both are non-stationary, implying that they follow a unit root process, regression estimates using these series in non-stationary form could lead to spurious relationships (Wooldridge 2008). There are several methods aimed at rendering such series stationary, with the most common being first differencing, or, in rare cases, multiple differences. According to Greene (2002), and Kripfganz and Schneider (2018), consistent estimates can be obtained from non-stationary variables in levels, provided the variables are cointegrated. A number of model specifications have been suggested to deal with non-stationary variables that are cointegrated including the autoregressive dynamic lag (ARDL) models, and the error correction model (ECM), using a maximum likelihood estimator (MLE).

We conduct two tests to check if COG and expected corn price, E(P_{Ct}), meet stationarity conditions, namely the augmented Dickey-Fuller (ADF), developed by Dickey and Fuller (1979) and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test as suggested by Kwiatkowski et al. (1992). Although both tests are concerned with stationarity, they have different null hypotheses. The null for the ADF is that the variable has a unit root, i.e., follows a random walk process (non-stationary), whereas the KPSS null is no unit root and thus a stationarity series. Applying both tests increases the confidence in test results, especially if both tests give the same conclusion.

The ADF null hypothesis is always H_0: unit root (non-stationary); however, the alternative differs and is predicated by whether the regression includes a drift term and whether the regression includes a constant term and time trend. Our regression has no time trend, but has a nonzero mean, and therefore it includes a constant term. A first-order autoregressive (AR) process can be represented by the expression

(3.10) $$y_t = \alpha + y_{t-1} + u_t$$

where u_t is defined as an independently and identically distributed (*i.i.d.*) error term, with mean zero. Under the Dickey-Fuller (DF) test, the following regression is fitted

(3.11) $$\Delta y_t = \alpha + \rho y_{t-1} + u_t$$

Given the ubiquity of serial correlation in time series, this model is likely to suffer from serial correlation and test results may be misleading without correction. The ADF test controls for possible serial correlation by modifying equation (3.11) and adding lagged differences leading to the following model:

(3.12) $$\Delta y_t = \alpha + \gamma y_{t-1} + \sum_{j=1}^{q} \varphi_j \Delta y_{t-j} + e_t$$

where e_t is *i.i.d*. with mean zero and constant variance and q is the number of optimal lags to be determined. Greene (2003) recommends using a testing down approach by sequentially evaluating the t statistic on the last coefficient until at the lag when the last coefficient is significant. Alternatively, one can use a combination of model fit and information criteria such as the Akaike information criteria (AIC) or the Bayesian information criteria (BIC). We use the AIC to select optimal lag length. The null for the ADF test H_0: $\gamma=0$, is equivalent to $\rho=1$ in equation (3.11) and implies that y_t follows a unit root process. The test statistic (Z_t) is $Z_t = \frac{\hat{\gamma}}{\hat{\sigma}_\gamma}$, where $\hat{\sigma}_\gamma$ is the standard error of $\hat{\gamma}$, but does not have a t-distribution.

Table 3.2 presents results of the ADF and KPSS tests. The ADF test fails to reject the null hypothesis of a unit root for COG and for corn price – implying both series have unit roots and are thus non-stationary. The KPSS test statistic was significant for both COG and corn price, and thus rejected the null hypothesis of stationarity. Results of the two tests complement each other, increasing the possibility that the series are indeed non-stationary. First differencing each series (COG and E(P_{Ct})) renders each series stationary, implying that the variables are I(1) non-stationary and that shocks to each of the two processes permanently affect the original variables. Following Greene (2002), if COG_t and P_{Ct} are both I(1), there may exist a coefficient (β) such that

$$\hat{\varepsilon}_t = COG_t - \beta P_{Ct} \qquad (3.13)$$

is I(0). That is, the two variables are each an I(1) process but their linear combination is an I(0) process. Two series that satisfy this condition are said to be cointegrated with a cointegrating vector [1, -β] (Greene 2002). A graphical inspection in Figure 1 indicates that the two series appear to move together, suggesting cointegration.

To test whether the series are cointegrated, we first perform a log transformation of COG and $P_{Ct,}$ because our empirical model is in log form. Next, we performed a two-step cointegration test suggested by Engle and Granger (1987), as well as Johansen's (1995) cointegration test.

Following Engle and Granger (1987), we regress log of COG on log of expected corn price via ordinary least squares (OLS) as expressed below:

(3.14) $$Ln(COG_t) = \alpha_0 + \alpha_1 Ln(P_{Ct}) + \zeta_t$$

where COG_t is cost of gain at time t, P_{Ct} is the expected corn price at time t, and ζ_t is the error term, with mean zero and constant variance. We then conduct a DF unit root test on the residuals, $\widehat{\zeta_t}$, as well as the Johansen's cointegration test. A Dickey-Fuller test rejects the presence of a unit root in the residuals at the 1% level, implying stationarity in the linear combination of log of COG and log of expected corn price, and hence cointegration. Similarly, Johansen's test fails to reject the null of one cointegrating equation (trace statistic = 3.318 against a critical value of 6.650). Both tests point to the presence of cointegration between *Ln* (*COG*) and *Ln* (P_{Ct}). Results of ADF, KPSS, and Johansen's cointegration tests are presented in Table 3.2. The broader implication is that COG and P_{Ct} drift considerably up and down, but do not deviate much from each other in the long-run, due to equilibrium forces. Because our data is cointegrated as evidenced by both graphical and statistical tests, it is appropriate to model the relationship in levels without making the data stationary.

Modeling and Estimation

The COG estimates reported in the data are averages of monthly closeout COG for the entire feeding period which is typically about five months in our data. Only one COG estimate per feeding period is reported. Such data structure reduces the sample, as we are unable to observe COG estimates for the other months during the feeding period. For example, a five-month change in COG can be calculated for the January to May period, for February to June, and likewise for March to July etc. In such instances, there is an overlap in COG estimates for a period of four months. Hansen and Hodrick (1980) show that such overlap in the data creates a moving average (MA) error term, making ordinary least squares (OLS) estimates inefficient and resulting hypothesis tests biased. Following Harri and Brorsen (2009) a basic model of COG and corn price can be expressed as

$$(3.15) \qquad y_t = \beta^{/} x_t + \eta_t$$

where y_t is the dependent variable, x_t is a vector of exogenous independent variables, and η_t is the error term, with E[η_t] = 0, E[η_t^2] $= \sigma_\eta^2$, and Cov [η_t , η_s] = 0 if $t \neq s$. Equation (3.15) represents the underlying data used to create the average sample. However, when data is average, the model can be formulated as

$$(3.16) \qquad Y_t = \beta^{/} X_t + \epsilon_t$$

where Y_t and X_t are averages of y_t and x_t respectively, averaged across a given time space – feeding period in this case. Data used to estimate (3.16) are formed by taking the average of the original monthly values of y_t and x_t, over the feeding period, K, as shown below:

$$(3.17) \qquad Y_t = \frac{1}{K}\sum_{j=1}^{K} y_j,\ X_t = \frac{1}{k}\sum_{j=1}^{K} x_j, \text{ and } \epsilon_t = \frac{1}{k}\sum_{j=1}^{K} \eta_t$$

Harri and Brorsen (2009) show that transforming the dependent and independent variables by summing over the overlapping periods creates an MA process error term. By this same argument, transformation of the dependent and independent variables by taking averages across overlapping periods should result in an MA process error term.

The theoretical model in equation (3.16) provides a guide for constructing the empirical model using average data like the Kansas feedlot data. Closeout COG is dependent on a number of factors, such as corn price, daily gain, placement weight and seasonality. Past research has shown that corn price helps explain much of the variation in COG relative to other factors, implying a strong relationship between corn price and COG (Langemeier 2015; Anderson and Trapp 2000a; Anderson and Trapp 2000b; Fox 1996; Albright et al. 1994; Mintert et al. 1991). Although placement weight appears in the theoretical model (2), it has very little variation in our data – likely due to its aggregated nature, and was thus insignificant in initial models. Following

Greene (2002), we conduct a likelihood ratio test of the model with and without placement weight to determine if the restriction is significant. Test results indicate that the restricted model (without placement weight) is not nested in the full model (with placement weight).

To construct a prediction model that is parsimonious and yet informative, we omit placement weight and model expected COG as a function of expected corn futures price, seasonality, and past COG. Empirically, we use a double log model so that we can interpret the coefficients on independent variables as percentage change in the dependent variable due to a one percentage change in a given independent variable. The empirical model has the following form:

(3.18) $$Ln(COG_{t+5}) = \beta_0 + \beta_1 Ln(COG_t) + \beta_2 Ln(P_{Ct}) + \beta_3 Ln(P_{Ct-1}) + \sum_{j=1}^{3} \delta_j S_{jt} + \varepsilon_t$$

where COG_{t+5} is closeout cost of gain at time t+5, for cattle placed at time t and expected to be finished in 5 months time, which is the COG realization at the end of the feeding period. COG_t is the closeout cost of gain for the most recently finished cattle. Thus, COG_t represents previous feeding period cost of gain. P_{Ct} is expected corn price at time t, P_{Ct-1} is the one time period lag of expected corn price, S_{jt} are seasonal dummy variables for placement season (winter, spring, and fall, with summer as the benchmark), and ε_t is MA process error term. Research involving Monte Carlo studies such as Harri and Brorsen (2009), Engle (1996), and Marcellino (1999) show that using OLS to estimate equation (2.16) – where aggregate data is used and a lagged dependent variable is one of the independent variables - could lead to biased estimates. However, consistent estimates can be obtained using maximum likelihood methods developed for time series models. We estimate the above model using a maximum likelihood estimator developed for ARIMA models in Stata version 14.2.

We estimated the COG predictor model in equation (2.16) using MLE to deal with the moving average error term stemming from the average nature of our data. We estimate three prediction models based on equation (2.16) varying the structure of the seasonality variables. The first model uses eleven monthly placement dummy variables; the second model uses the four seasons (i.e., spring, summer, fall, and winter); and the third model uses sinusoidal functions (sine cosine transformation) of placement month. We also estimate a naïve model equation (2.17), by regressing closeout COG on its previous period values (COG from 5 months ago), without controlling for seasonality.

Prediction Accuracy

We compare the model's statistical prediction properties with those of the naïve model currently used to predict closeout COG in order to evaluate the prediction accuracy of the constructed model. We test both in-sample and out-of-sample. Under the naïve model, closeout COG for the feeding period just finished is used as a predictor of current month's placement COG. Thus, the naïve model can be expressed by the equation

(3.19) $$COG_{t+5} = \alpha_0 + \alpha_1 COG_t + \nu_t.$$

For out-of-sample prediction accuracy, we split the sample into two parts. One part spanning January 1992 – December 2008 (204 observations) is used to estimate the model while the other part spanning January 2009 – June 2017 (102 observations) is held out for out-of-sample prediction.

Following Fair (1986), the three most common measures of predictive accuracy for statistical forecasting models are root mean squared error (RMSE), mean absolute error (MAE), and Thiel's inequality coefficient *U*. Both RMSE and MAE are scale dependent and thus can only be used to measure forecast or prediction performance of the same series across different models. Another common measure is the mean absolute percentage error (MAPE), which measures the prediction accuracy of series with different scales of the dependent variable across different models. Hence it is not scale dependent. For example, MAPE can be used to compare a model with a log transformed dependent variable versus a model with dependent variable in levels. Smaller values of each of these statistics indicate higher prediction accuracy. RMSE and MAPE are defined as

$$(3.20) \qquad RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(Y_t^p - Y_t^a\right)^2}$$

$$(3.21) \qquad MAPE = \left(\frac{100}{T}\sum_{t=1}^{T}\left|\frac{Y_t^p - Y_t^a}{Y_t^a}\right|\right)$$

where Y_t^p is the predicted value of the dependent variable, Y_t^a is the actual reported value of the dependent variable, and T is the sample size.

Another useful measure of prediction accuracy is Theil's inequality coefficient (U), defined as

$$(3.22) \qquad U = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(Y_t^p - Y_t^a\right)^2}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(Y_t^p\right)^2} + \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(Y_t^a\right)^2}}$$

The numerator of U is RMSE and the scaling of the denominator is such that U ranges between 0 and 1, with 0 representing a perfect prediction and 1 corresponding to poor predictive performance (Pindyck and Rubinfeld 1998).

The above measures are based on calculation of error, mean squared error (MSE). As a result, it is highly likely that estimation of prediction accuracy across all three measures will differ for each model. One important question to ask then, is how likely is it that the differences observed are due to chance? This question is answered by determining if the observed accuracy difference between the prediction model and the naïve model is statistically significant. Performance of the two models can be compared using the Diebold Mariano (DM), a parametric test suggested by Diebold and Mariano (1995). The DM tests the hypothesis that there is no difference between two competing prediction models. Given two prediction models, 1 and 2, we define the prediction error as $e_{i,t} = \hat{Y}_{i,t} - \hat{Y}_t, \quad i = 1, 2$ and the respective loss functions as $g(e_{1,t})$ and $g(e_{2,t})$, both of which are assumed to be linear. Following Diebold and Mariano (1995), we define a loss function differential between the two prediction models as $d_t = g\left(e_{1,t}\right) - g(e_{2,t})$, and conclude that the two models have equal accuracy if and only if the loss differential has zero expectation for all t, resulting in the following null hypothesis: H_0: $E(d_t) = g(e_{1,t}) - g(e_{2,t}) = 0$. The test statistic derived under the DM test has an asymptotic standard normal distribution (Kim and Brorsen 2017) and is thus valid in large samples like ours.

Results and Discussion

Tables 3.4, 3.5, and 3.6 present estimation results of the three constructed models, which vary by seasonality variables, and Table 3.3 presents estimates of the naïve model. With regards to the three constructed models, the difference in estimates across all three models were small, although the second model outperformed the other two in terms of prediction accuracy. In addition, we performed a log-likelihood ratio test comparing the model which uses monthly dummies and the one using the four seasonal dummies (winter, spring, summer, and fall). Results of the test show that the restriction imposed by the four seasonal dummies relative to monthly dummies could not be rejected. This suggests that the model with the four seasonal dummy variables is adequate. Besides predicting expected COG with higher accuracy, the model with four seasonal dummy variables is also parsimonious. Thus, the discussion in this section is based on the second model (Table 3.5). Results in Table 3.5 indicate that previous closeout COG prior to current placement, and expected corn price relevant to the feeding period are positive and statistically significant predictors of COG. In terms of magnitude, a one percent increase in previous closeout COG will lead to a 0.45 percentage point increase in expected COG for current month's placement, holding all else constant.

In regards to expected corn price, a percentage increase in the contemporaneous expected corn price increases predicted COG by 0.18 percent. In terms of lagged expected corn price, a one percent increase in lagged expected corn price leads to a 0.23 percent increase in predicted COG. Total expected corn price effect on COG is 0.41, achieved by adding the two expected corn price effects (0.18+0.23). Our estimate of the effect of expected corn price effect on COG is lower compared to others who have attempted to model this relationship (e.g., Anderson and Trapp 2000a; Maday 1996; Fox 1996). Albright et al. (1994) find that, on average, corn price explains 64 percent of the variation in COG. Fox (1996) and Maday (1996) assume a one-to-one relationship between corn price and COG. Anderson and Trapp (2000a) provide a more comprehensive modeling approach and find a 0.53 percent increase in COG for a percentage increase in corn price. This discrepancy is most likely due to the differences in model specifications. Anderson and Trapp (2000a) acknowledge the influence of past COG and attempt to account for it by including past corn prices which feedlots reported to have bought the corn at as proxies for past COG. Our model on the other hand explicitly models for COG dynamics by including lagged COG values. Hence, the effect of corn price in our model is dampened by the COG dynamics.

Theoretically, the dynamics of COG can be explained by the fact that feedlots use information from previous feeding periods to make adjustments to future feeding management. As noted by Belasco et al. (2009), Anderson and Trapp (2000b) and Marsh (1995), cattle feeders will tend to substitute other feed including forage as corn price varies. In addition, changes in corn price will likely induce changes in placement weight by feedlots, which in turn affects feed conversion efficiency. Ultimately, changes in feed conversion efficiency as a result of changes in placement weight will affect COG. However, such adjustments may not be immediate, because feedlots need to adjust their feeding facilities to accommodate bigger animals, thus creating a lag in time between increasing COG and feedlots implementing cost saving measures. Generally, cattle have gotten bigger over-time, and this could affect the results of this study because larger cattle tend to have a higher COG per pound than smaller cattle. This could lead to upward bias of Cog prediction. However, the increase in cattle size is flanked by improvements in feeding efficiency and management practices. Thus to the extent that these improvements help reduce COG, our results may adjust for increase in cattle size.

Anderson and Trapp (2000) using sinusoidal seasonal variables find that seasonality significantly influences COG. However, the use of sinusoidal variables does not clearly illustrate how the different seasons influence cost of gain. To account for seasonal effects, we included three dummy variables, each representing a placement season (winter, spring, and fall), with summer as the benchmark season. Results indicate lower closeout COG for winter and spring placements relative to summer, whereas fall placements had higher COG, albeit insignificant. Cattle placed in the winter and spring months will finish in the warmer months (summer and fall) when the temperatures are warm enough for efficient feeding. This finding corresponds fairly well with Belasco, Ghosh and Goodwin (2009) who find that spring placements had significantly lower feed conversion rates (higher efficiency), implying lower COG, relative to summer placements. However, they find winter placements not to be statistically significantly different from summer (although with a negative sign, which is similar to ours). Further, Albright et al. (1994) note that placements between February and August tend to have below average COG.

Table 3.7 presents a summary of prediction accuracy measures for the in-sample and out-of-sample prediction of constructed model (CM) and the naïve model (NM). The CM outperformed the NM in terms of both in-sample and out-of-sample prediction. In terms of in-sample prediction, the CM outperformed the NM on both stand-alone measures (RMSE and MAPE) as well as on the relative measure, Theil's inequality coefficient (*U*). With RMSE of 0.029 and MAPE of 0.006, the constructed model outperforms the naïve model, which had RMSE of about 0.1 and MAPE of 0.02. In terms of Theil's *U* test, the constructed model had a coefficient of 0.83 (less than 1), whereas the naïve model had a coefficient of 3.2. Further, the DM test of accuracy equality between the naïve and constructed model was rejected, implying that the difference in accuracy between the two models is not zero. This reinforces findings from the other three tests. With regards to out-of-sample, the DM tests based on MSE and MAE both indicate that the CM predicts COG more accurately than NM, and that the difference in prediction error between the two models is not zero. Figure 3.2 and 3.3 show an in-sample comparison of the naïve model versus reported COG and ex-ante prediction from constructed model versus reported COG, respectively. The constructed model maps the reported COG with more accuracy than the naïve model.

Implication for Breakeven Analysis

Given that the constructed model outperforms the naïve model, breakeven price estimates based on COG predicted from the NM is expected to be less accurate relative to that obtained based on COG predicted from the CM. Hence, the price spread as calculated using equation (3.7) is expected to be smaller if the breakeven is calculated based on COG predicted from the CM, compared to the breakeven estimated based on COG obtained using the NM. This is primarily because the constructed model accounts for the variation in expected price of corn, a major ingredient in feedlot feed rations. The naïve model on the other hand is a function of previous reported closeout COG, and is thus unable to capture variation in expected corn price. Figures 3.4 – 3.8 show the price spreads based on average prices from 2000 – 2016. When the CM predicted COG is used to calculate breakeven price, results show decreasing spreads between feeder market price and breakeven price in the first half of the year, whereas in the second half of year, the spread remains very similar to that from the NM. An exception is the 600-650 pound weight category, where in the second half of the year, the spread from the constructed actually increases substantially relative to that of the naïve model (Figure 3.5). In addition to graphical analysis, we further analyzed the spreads by month averaged across all weight categories using mean absolute deviation (MAD). Results in Table 3.8 indicate that overall, the spread between 800-850 pound and lighter feeders based on constructed model COG prediction are larger than those based on naïve model COG prediction. A look at monthly spreads reveals that in late winter and spring, spreads between 800-850 pound and lighter feeders based on constructed COG prediction are smaller relative to those based on naïve model prediction. However, the reverse is true in the second half of the year summer and fall), where the naïve model spreads tend to be lower than spreads from constructed model.

This is an indication that besides COG, other market conditions may be influencing the spread between feeder market price and feedlot breakeven price. Peel and Riley (2018) point out that feeder cattle price fluctuations tend to be predicated on broad economic conditions over time, such as feeder cattle demand and supply, which obviously cannot be captured by a model like ours. Thus, the most accurate ex-ante COG prediction will not explain all of the price spread detected in the data overtime. Further, given the small changes in expected corn price between feeding period closeouts, the relative difference in estimated COG from the constructed model and naïve model is also small and so there is little difference in breakeven estimates from the two models. However, in the event of significant exogenous shocks to corn production (e.g., drought-reduced yield), the constructed model will capture the effects of such a shock via the corn futures market, which will be signaled by a wide upward swing in corn futures price. Since our model includes corn futures market prices relevant to the feeding period, the increase in corn futures price will translate into increased expected COG by a larger margin. Comparatively, the naïve model will only capture such a change with a lag, since it depends on previous closeout COG. Under such circumstances, the improvement in COG prediction by the CM will be more pronounced.

Conclusion

This study set out to address two objectives. First was to construct a COG prediction model, and second was to use the predicted COG to estimate feedlot purchase breakeven price for lighter feeders relative to heavier feeders (assumed to be 800-850 pounds). An ex-ante feedlot cost of gain model, which accounts for the overlapping nature of feedlot placement and closeout data, was constructed and statistically tested against the naïve model typically used in feedlot cost of gain prediction. The constructed model uses expected corn price representative of the entire feeding period and also accounts for the fact that feedlots typically procure corn through forward contracting and keep inventories for future use. Statistical evaluation of model prediction performance indicates that the constructed model outperforms the naïve model in terms of predicting feedlot cost of gain. The naïve model assumes future closeout COG to be a function of current closeout COG only. The constructed model results indicate that in addition to current closeout COG, expected corn price and its lagged values, as well as seasonality are important predictors of closeout COG. Regarding seasonality, placements earlier in the year (winter and spring) tend to be associated with lower COG relative to later placements.

The constructed COG prediction model is used to predict future COG and to calculate feedlot breakeven price for various feeder weights relative to 800-850 pound steers. Results show a reduction in spreads between 800-850 pound feeder market price and breakeven price for lighter weights mostly in the first half of the year, albeit small compared to breakeven prices calculated using the naïve COG expectation. This finding could be an indication that more accurate estimation of feedlot breakeven price is not only a function of COG, but also of other market conditions not captured by COG, such as changes in feeder cattle supply and variations in demand across weight and time. If COG completely explained the difference between feeder market price and feedlot breakeven price, then our model would have significantly reduced the spread between feeder market price and feedlot breakeven price and its apparent seasonality. Thus, any remaining difference between estimated breakeven price and feeder market price can be attributed to market pricing inefficiency of lighter feeder cattle or other market conditions. Although the constructed model may not capture broader market conditions, it does account for expected changes in corn demand and supply as reflected in the corn futures price. Given the significance of corn in feedlot rations, it is reasonable to expect the COG prediction from the constructed model to be more representative of the COG - corn price relationship, which tends to vary over-time. This is in comparison to the naïve expectation which is based solely on previous COG estimates.

As this study has shown, accurate prediction of COG is only a small part in explaining the spreads between steer feeder market price and feedlot breakeven price for lighter weights, when corn prices fluctuate only marginally. The other part could be attributed to other factors including availability of grass during spring, which may affect COG in the sense that it reduces supply of lighter feeders to feedlots, because backgrounders would rather keep animals a little longer on grass before passing them on to feedlots. Thus, most of the cattle supplied to feedlots in late spring/summer will be heavier than in the fall and winter, when there is less grass available. This affects feedlot COG since it is mostly heavier feeders being placed on feed, thus increasing COG per pound.

One of the significant limitations of this study is the use of aggregate data which limits the functional form to linear. However, it is possible that the COG model follows a nonlinear function. Thus, with availability of pen level or more disaggregate data, future studies should consider other functional forms. Our model is the inability to capture changes in supply and demand of feeder cattle across time.

References

Albright, M.L., T.C. Schroeder, and M.R. Langemeier. 1994. “Determinants of Cattle Feeding Cost of Gain Variability.” *Journal of Production Agriculture* 7:206–210.

Anderson, J.D., and J.N. Trapp. 2000a. “Corn Price Effects on Cost of Gain for Feedlot Cattle : Implications for Breakeven Budgeting.” *Journal of Agricultural and Resource Economics* 25(2):669–679.

Anderson, J.D., and J.N. Trapp. 2000b. “The Dynamics of Feeder Cattle Market Responses to Corn Price Change.” *Journal of Agricultural and Applied Economics* 3(December):493–505.

Antonovitz, F., and R. Green. 1990. “Alternative Estimates of Fed Beef Supply Response to Risk.” *American Journal of Agricultural Economics* 72(2):475–487.

Belasco, E.J., S.K. Ghosh, and B.K. Goodwin. 2009. “A Multivariate Evaluation of *Ex ante* Risks Associated with Fed Cattle Production.” *American Journal of Agricultural Economics* 91(2):431–443.

Belasco, E.J., M.R. Taylor, B.K. Goodwin, and T.C. Schroeder. 2009. “Probabilistic Models of Yield, Price, and Revenue Risks for Fed Cattle Production.” *Journal of Agricultural and Applied Economics* 41(01):91–105.

Dhuyvetter, K.C., and T.C. Schroeder. 2000. “Price-Weight Relationships for Feeder Cattle.” *Canadian Journal of Agricultural Economics* 48(3):299–310.

Dickey, D.A., and W.A. Fuller. 1979. “Distribution of the Estimators for Autoregressive Time Series with a Unit Root.” *Journal of the American Statistical Association* 74(366a):427–431.

Diebold, F.X., and R.S. Mariano. 1995. “Comparing Predictive Accuracy.” *Journal of Business and Economic Statistics* 13(3):253–263.

Eales, J.S., B.K. Engel, R.J. Hauser, and S.R. Thompson. 1990. “Grain Price Expextations of Illinois Farmers and Grain Merchandisers.” *American Journal of Agricultural Economics* 72(3):701–708.

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

Fox, B. 1996. “Why Worry About $5 Corn? .” *National Cattleman* 11:24-27.

Gardner, B. 1976. “Futures Prices in Supply Analysis.” *American Journal of Agricultural Economics* 58(1):81–84.

Greene, W. 2002. *Econometric Analysis* 5th ed. New Jersey: Prentice Hall.

Hansen, L.P., and R.J. Hodrick. 1980. “Forward Exchange Rates as Optimal Predictors of Future

Spot Rates : An Econometric Analysis." *Journal of Political Economy* 88(5):829–853.

Harri, A., and B.W. Brorsen. 2009. "The Overlapping Data Problem." *Quantitative and Qualitative Analysis in Social Sciences* 3(3):78–115.

Kim, S.W., and B.W. Brorsen. 2017. "Forecasting Urea Prices." *Applied Economics* 49(49):4970–4981.

Kansas State University Agricultural Experiment Station and Cooperative Extension Service. 2016. "Kansas Feedlot Performance and Feed Cost Summary 2016 Annual Review." *Focus on Feedlots*. Available at: www.asi.k-state.edu/about/newsletters/focus-on-feedlots/.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, and Y. Shin. 1992. "Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root." *Journal of Econometrics* 54:159–178.

Langemeier, M. 2015. "Feeding Cost of Gain and Net Returns for Cattle Finishing." *farmdoc daily* 5:190. Available at: https://ageconsearch.umn.edu/record/229544/files/fdd141015.pdf

Langemeier, M., T. Schroeder, and J. Mintert. 1992. "Determinants of Cattle Finishing Profitability." *Journal of Agricultural and Applied Economics* 24(2):41–47.

MacDonald, A., and T. Schroeder. 2003. "Fed Cattle Profit Determinants Under Grid Pricing." *Journal of Agricultural and Applied Economics* 35(1):97–106.

Maday, J. 1996. "Living in a Weather Market." *Drover's Journal* 124:24–26.

Mark, D.R., and T.C. Schroeder. 2002. "Effects of Weather on Average Daily Gain and Profitability." *Kansas Agricultural Experiment Station Research Reports*: Vol. 0: Iss. 1. Avaialable at: https://doi.org/10.4148/2378-5977.1748

Mark, D.R., T.C. Schroeder, and R. Jones. 2000. "Identifying Economic Risk in Cattle Feeding." *Journal of Agribusiness* 18(3):331–344.

Marsh, J.M. 1999. "Economic Factors Determining Changes in Dressed Weights of Live Cattle and Hogs." *Journal of Agricultural and Resource Economics* 24(2):313–326.

Mintert, J.R, T.C. Schroeder, M.R. Langemeier, and M.L. Albright. 1994. "Factors Affecting Cattle Finishing Profitability." *Kansas State University, Agricultural Experiment Station and Cooperative*.

Peel, D.S., and J.M. Riley. 2018. "Feeder Cattle Price Fundamentals." *Western Economic Forum* 16(2):6–19.

Wooldridge, J.M. 2008. *Introductory Econometrics: A Modern Approach*, Mason, OH: South-Western College Publishing.

Zhao, H., X. Du, and D.A. Hennessy. 2011. "Pass-Through in United States Beef Cattle Prices: A Test of Ricardian Rent Theory." *Empirical Economics* 40(2):497–508.

Table 3.1. Summary Statistics for Steer Closeout Cost of Gain, Expected Corn Price, and Days on Feed January 1992 – June 2017

Variable	N	Mean	SD	Min	Max
Placement weight (lbs)	306	784.59	40.87	681.00	877.00
Average cost of gain ($/Cwt	306	67.75	20.84	42.04	133.72
Expected corn price	306	3.41	1.45	1.89	7.63
Average days of feed	306	150.97	12.10	119.00	186.00

Source: Livestock Marketing Information Center, 2017

Table 3.2. Stationarity and Cointegration Tests

Test	Variable	ADF Test Statistic	KPSS Test Statistic	Johansen's Trace Statistic	5% Critical Value
ADF/KPSS	Steer cost of gain (*COG*)	-1.146	0.257**		
	Expected corn price (P_C)	-1.305	0.213*		
	$\hat{\zeta}_t$	-4.025**			
Johansen's (H_0: r ≤1 H_A: r=2)				3.318	6.650

Note: * p<0.05; ** p<0.01

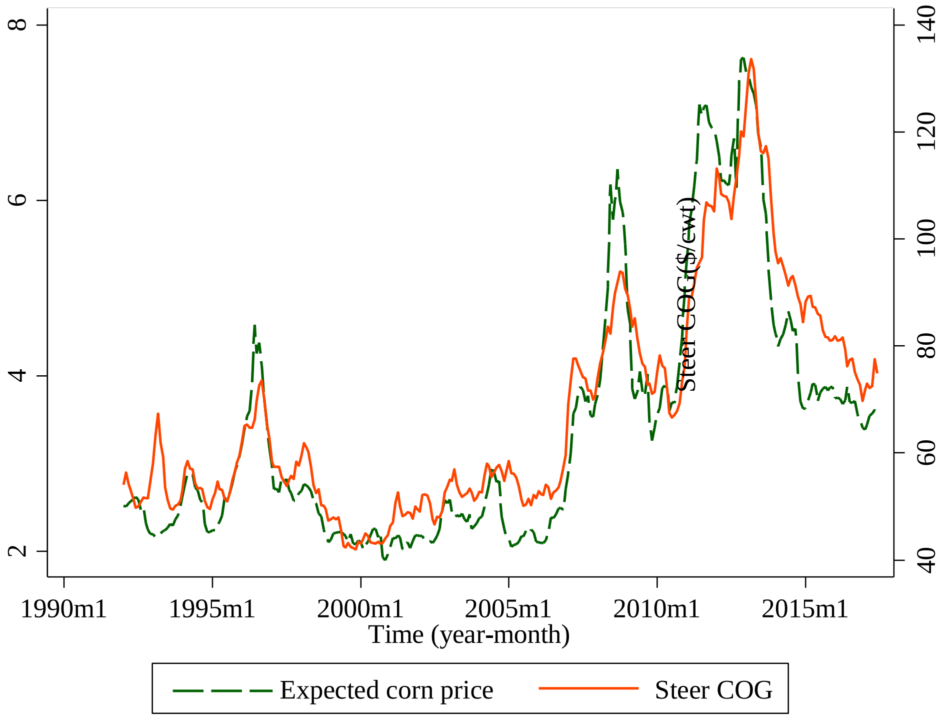


Figure 3.1. Steer Cost of Gain and Expected Corn Price Series – January 1992 – June 2017

Table 3.3. Naïve Prediction Model of Expected Closeout Cost of Gain

Variable	Coef.	Std. Err.	[95% Conf. Interval	
Intercept	4.784	1.340**	2.149	7.419
Current cost of gain	0.934	0.019**	0.896	0.971

Note: + p<0.1; * p<0.05; ** p<0.01

Table 3.4. Constructed Prediction Model (Equation 18) of Log Expected Closeout Cost of Gain with Monthly Dummy Variables

Variable	**Coef.**	**Std. Err**	**[95% Conf.Interval]**	
Intercept	1.847	0.160**	1.534	2.160
Log current cost of gain	0.446	0.047**	0.354	0.537
Log expected corn price	0.193	0.029**	0.136	0.250
Log previous corn price	0.22	0.032**	0.158	0.282
January placement	-0.039	0.006**	-0.051	-0.028
February placement	-0.049	0.007**	-0.063	-0.034
March placement	-0.04	0.008**	-0.057	-0.024
April placement	-0.027	0.010**	-0.046	-0.008
May placement	-0.03	0.010**	-0.049	-0.011
June placement	-0.016	0.010+	-0.035	0.003
July placement	0.002	0.009	-0.016	0.020
August placement	-0.006	0.009	-0.024	0.013
September placement	0.007	0.009	-0.01	0.024
October placement	0.016	0.008*	0.001	0.031
November placement	0.022	0.006**	0.011	0.032
Moving Average				
L.ma	1.025	0.069**	0.89	1.161
L2.ma	0.916	0.098**	0.724	1.108
L3.ma	0.972	0.103**	0.771	1.173
L4.ma	0.899	0.107**	0.688	1.109
L5.ma	0.265	0.096**	0.077	0.452
L6.ma	0.06	0.066	-0.07	0.190

Note: + p<0.1; * p<0.05; ** p<0.01; L.ma - L6.ma are the moving average terms

Table 3.5. Constructed Prediction Model (Equation 16) of Log Expected Closeout Cost of Gain with Seasonal Variables

Variable	**Coef.**	**Std. Err.**	**[95% Conf. Interval**	
Intercept	1.837	0.189**	1.467	2.207
Log current cost of gain	0.450	0.055**	0.342	0.559
Log expected corn price	0.171	0.029**	0.115	0.227
Log previous period corn price	0.229	0.032**	0.167	0.291
Winter placement	-0.013	0.006*	-0.025	-0.001
Spring placement	-0.011	0.005*	-0.021	-0.001
Fall placement	0.005	0.004	-0.002	0.013
Moving Average				
L.ma	1.135	0.069**	1.000	1.269
L2.ma	1.002	0.088**	0.829	1.174
L3.ma	0.960	0.094**	0.776	1.145
L4.ma	0.849	0.098**	0.657	1.040
L5.ma	0.135	0.071+	-0.005	0.275

Note: + $p<0.1$; * $p<0.05$; ** $p<0.01$; L.ma – L5.ma are the moving average terms

Table 3.6. Constructed Prediction Model (Equation 16) of Log Expected Closeout Cost of Gain with Sinusoidal Seasonal Variables

Variable	**Coef.**	**Std. Err**	**[95% Conf.Interval]**	
Intercept	1.837	0.179**	1.487	2.187
Log current cost of gain	0.445	0.052**	0.343	0.548
Log expected corn price	0.175	0.031**	0.116	0.235
Log previous period corn price	0.237	0.033**	0.172	0.302
Cosine 12 months	0.005	0.006	-0.007	0.017
Sine 12 months	-0.031	0.007**	-0.044	-0.018
Moving Average				
L.ma	1.079	0.073**	0.937	1.221
L2.ma	0.938	0.103**	0.737	1.139
L3.ma	0.877	0.123**	0.637	1.118
L4.ma	0.731	0.124**	0.488	0.973
L5.ma	0.131	0.071+	-0.008	0.271

Note: + p<0.1; * p<0.05; ** p<0.01; L.ma – L5.ma are the moving average terms

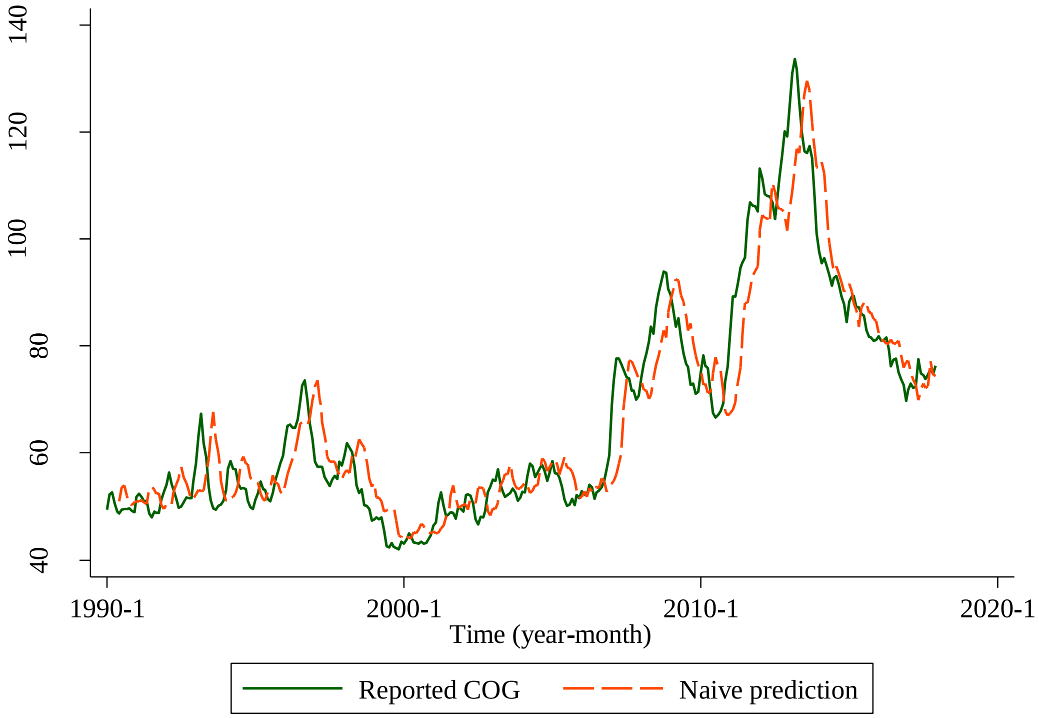


Figure 3.2. Steer Cost of Gain from Naïve Model Prediction and Reported Cost of Gain

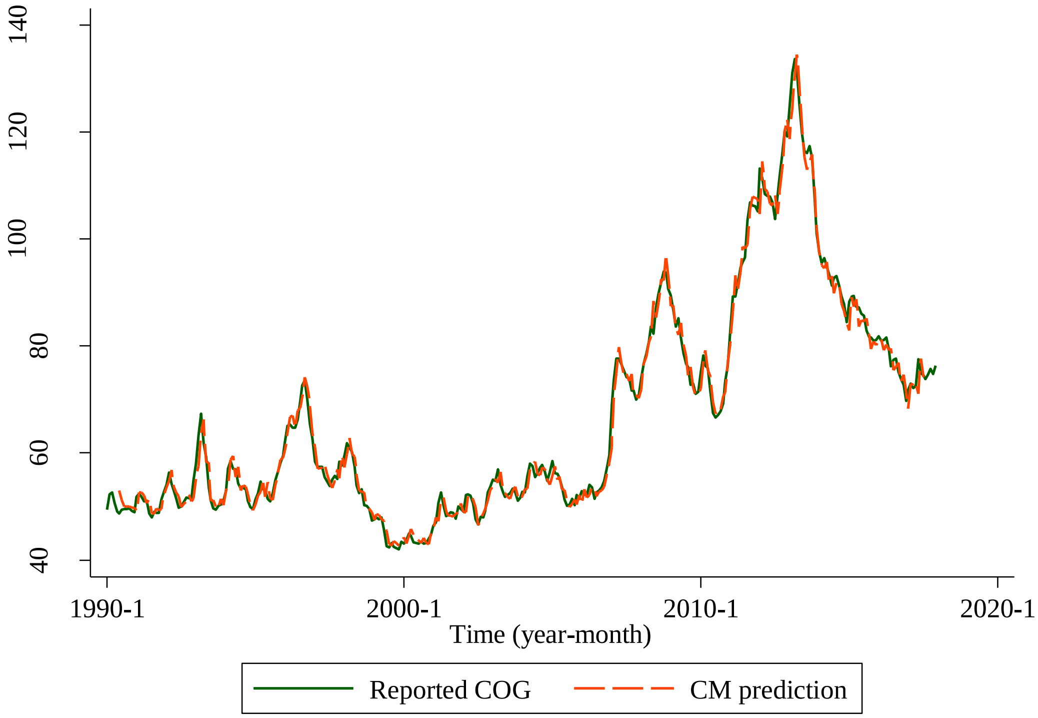


Figure 3.3. Comparison of Steer Cost of Gain from Constructed Model Prediction and from Reported Cost of Gain

Table 3.7. In-Sample and Out-of-Sample Prediction Accuracy Measures for Constructed Model and Naïve Model

Test Statistic	Constructed Model (CM)	Naïve Model (NM)
RMSE	0.029	0.099
MAPE	0.006	0.019
Theil's U	0.832	3.015
DM (MAE)	0.023	0.076
Out-of-sample DM (MSE)[6]	6.652	69.950
Out-of-sample DM (MAE)	1.919	6.095

[6] Out-of-sample predictions are made for time period spanning January 2009 – June 2017

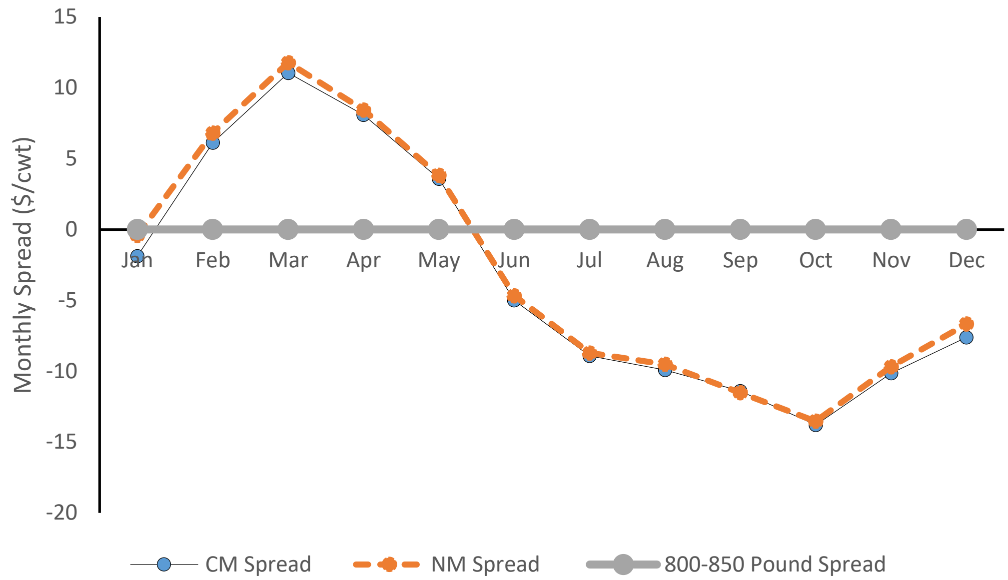


Figure 3.4. Spread Between Breakeven Prices and Market Prices for 550-600 Pound Feeders (Steers). Breakeven Prices are Relative to 800-850 Pound Feeders.

Note: CM and NM are constructed model and naïve model respectively.

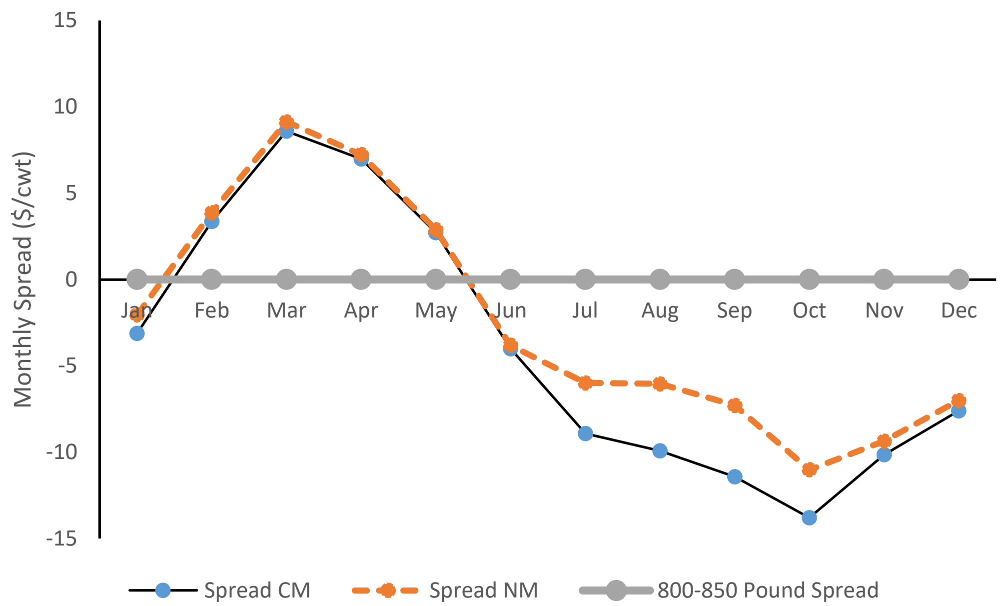


Figure 3.5. Spread Between Breakeven Prices and Market Prices for 600-650 Pound Feeders (Steers). Breakeven Prices are Relative to 800-850 Pound Feeders.

Note: CM and NM are constructed model and naïve model respectively

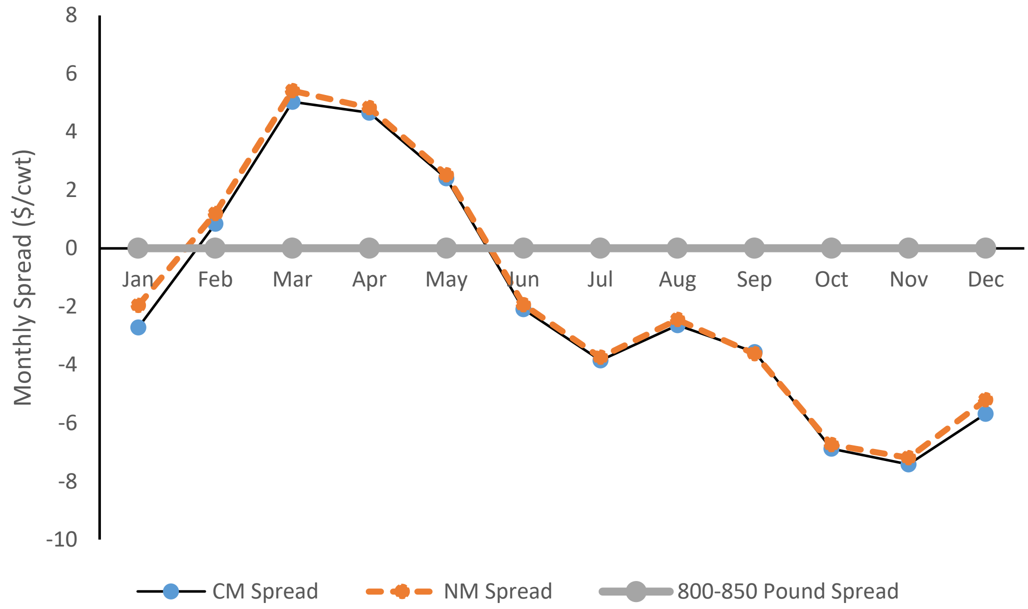


Figure 3.6. Spread Between Breakeven Prices and Market Prices for 650-700 Pound Feeders (Steers). Breakeven Prices are Relative to 800-850 Pound Feeders.

Note: CM and NM are constructed model and naïve model respectively

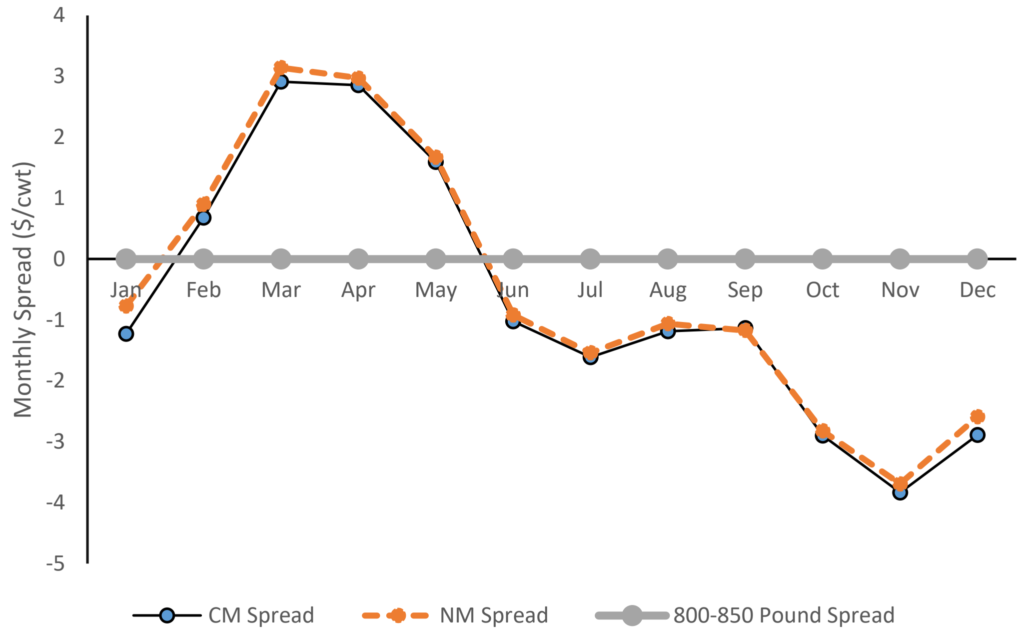


Figure 3.7. Spread Between Breakeven Prices and Market Prices for 700-750 Pound Feeders (Steers). Breakeven Prices are Relative to 800-850 Pound Feeders.

Note: CM and NM are constructed model and naïve model respectively

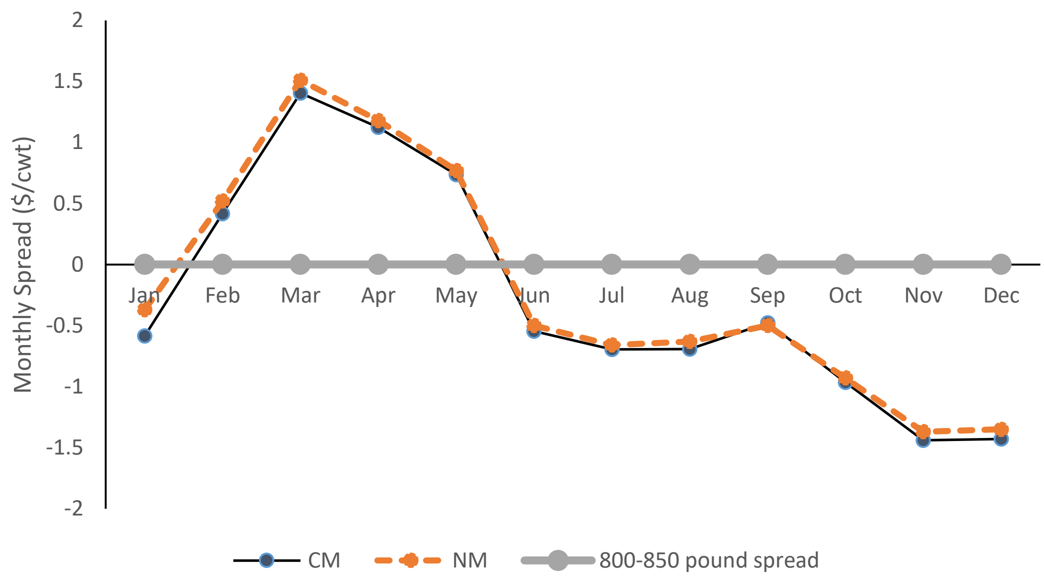


Figure 3.8. Spread Between Breakeven Prices and Market Prices for 750-800 Pound Feeders (Steers). Breakeven Prices are Relative to 800-850 Pound Feeders.

Note: CM and NM are constructed model and naïve model respectively

Table 3.8. Monthly Spreads Between Breakeven Prices and Market Prices ($/cwt).

Month	MAD CM	MAD NM
January	1.91	1.11
February	2.28	2.66
March	5.80	6.19
April	4.74	4.93
May	2.21	2.34
June	2.53	2.37
July	4.80	4.13
August	4.87	3.94
September	5.60	4.82
October	7.66	7.01
November	6.59	6.26
December	5.04	4.56
Average	4.50	4.19

CHAPTER IV

JOINT ADOPTION OF COW-CALF PRODCUTION PRACTICES

Introduction

Value enhancement for beef calves at marketing remains a major focus area of livestock research and extension efforts. Basic value-added beef production and management practices such as castration, dehorning, and vaccination help improve animal health, meat quality, and weight gain in later stages of production (Williams et al. 2013; Schumacher, Schroeder and Tonsor 2012). For cattle feeders (calf buyers) improved animal health due to preconditioning implies lower medical treatment costs and enhanced weight gain, which when taken together lead to reduced cost of production. Thus cattle feeders have an incentive to pay a premium for preconditioned calves. Empirical studies (Ward, Powell and Gadberry 2019; Williams et al. 2013; Dhuyvetter et al. 2005) indicate that value-added beef production practices attract price premiums and could improve producer profits. A study by (Behrends, Field and Conway 2001) based on a survey of feedlot managers indicates positive willingness to pay premiums for numerous value-added practices.

Despite the documented benefits of value-added practices, most producers have not adopted them (Schumacher, Raper, and Peel 2017).

This disjoint between expected benefits of adoption and lower than expected adoption rates has given rise to a number of studies (e.g., Schumacher, Raper, and Peel 2017; Williams et al. 2013; Pruitt et al. 2012) aimed at better understanding practice adoption. Generally, these studies find producer characteristics such as age, level of education, participation in producer education programs, income, and percent of cattle income to total income influence adoption decisions. Other important determinants of practice adoption include herd size and geographical location.

Existing studies pertaining to adoption of value-added production practices tend to treat practices individually and by implication ignore the possibility that some practices are more likely to be jointly adopted, such as castration and dehorning. For example, Williams et al. (2013) model determinants of number of practices adopted using a negative binomial model but do not examine joint adoption of specific practices. Popp et al. (1999) and Ward et al. (2008) use logit models to explain determinants of adoption for individual practices. Ward et al. (2008) disaggregate producers by scale of operation and level of dependence on cattle income, but still treat practices individually. Hence, these studies impose independence among practices by treating them individually. Some practices, for example, are typically adopted jointly; hence, adoption decisions across such practices are correlated. Treating adoption of individual practices as independent decisions leads to loss of information because correlation information is implicitly disregarded.

This study examines correlation among practices by clustering (bundling) practices based on the likelihood of joint adoption, thus helping provide a reference point upon which adoption decision modeling studies can build. Further, given that data used in most studies is becoming dated, a study using recent data to update understanding of the subject is warranted. This study uses data from a 2018 survey of Oklahoma cow-calf producers, providing an updated understanding of producer practice adoption. This information could be useful in improving targeting and design of extension programs.

Against this backdrop, this study sets out to address two objectives; 1) identify and construct bundles of practices likely to be jointly adopted with their associated probabilities; and 2) estimate the probability of adopting a practice conditional on adopting a given bundle.

Methods

At the core of this study is the need to identify practices that producers are more likely to adopt as a bundle. Toward that end, we apply the association rule learning suggested by Hahsler, Grün and Hornik (2005) and, specifically, the market basket analysis (MBA). Market basket analysis is an association rule learning technique primarily applied in retail marketing to identify sets of items (i.e., item sets) that customers tend to purchase together and the associated probability of doing so (Linoff and Berry 2011; Shaw et al. 2001; Agrawal, Road and Jose 1993). In addition to identifying items that are purchased together, MBA permits identification of items that are likely to also be purchased conditional on a customer's purchase of a given item set, with a given probability. This is based on the theory that if a customer purchases a certain item set, they are more (or less) likely to purchase another specific item or group of items (e.g., see Shaw et al. 2001; Chen et al. 2005). Here we define items as practices adopted by a producer. Similar to the item set, we define practice bundles as practices more likely to be jointly adopted by a producer. The item set (practice bundle) is referred to as antecedent and the additional item (practice) likely to be purchased (adopted) conditional on purchasing (adopting) the item set (practice bundle) is referred to as a consequent.

Following Agrawal et al.'s (1993) original formulation of association rule, let $I = \{i_1, i_2, i_3 \ldots, i_7\}$ be a set of 7 practices a producer is expected to adopt with each practice represented as a binary variable where 1 represents adoption and 0 represents non-adoption. Further, let T be the number of observations for which we have information regarding practice adoption among producers that adopted at least one of the 7 practices. Each observation of the sample (T) comprises all or a subset of practices in I. Each entity in T thus contains a bundle of practices whose size can range from 1-7.

The association rule can thus be defined by the following formulation $X \Longrightarrow Y$, where $X \subset I$, $Y \subset I$, and $X \cap Y = 0$. With this formulation, it is possible to identify, with a certain probability, a practice bundle *X* most likely to be adopted by a producer, and the practice *Y* likely to be adopted conditional on adopting bundle *X*. Each rule contains two different practice bundles, *X* and *Y*- with *Y* comprising only one practice - where *X* is referred to as the antecedent or left-hand side (LHS) and *Y* is the consequent or right-hand side (RHS). Following Hahsler et al. (2005), among the important parameters to consider in evaluating precision of discovered association rules are support, confidence, and lift. Support measures the frequency with which an item set, or practice bundle in this case, appears with one other practice in the dataset. Support is thus defined as proportion of observations out of total observations T, which contains the practice bundle *X* and one other practice Y. Practice bundles that are more frequent in the data will have higher support than those less frequently adopted. Mathematically, support can be expressed by the equation

(4.1) $$support\ (X \cup Y) = \frac{Frequency\ (X \cup Y)}{T}$$

We can also define support for bundle *X* as $support(X) = \frac{Frequency\ (X)}{T}$ and support for *Y* as $support(Y) = \frac{Frequency(Y)}{T}$.

Setting a minimum support threshold is helpful in identifying practice bundles that occur frequently enough in the data to warrant further analysis. Bundles with support below the minimum threshold are considered not to have enough information about the practices contained, suggesting that no meaningful conclusions can be inferred from that particular rule. Confidence is a measure of how often a rule holds. Specifically, it measures the percentage of observations in which *Y* has been adopted given that bundle *X* is already adopted, i.e. the conditional probability of *Y* given *X* (e.g., see Hipp, Güntzer and Nakhaeizadeh 2007).

Mathematically, confidence can be expressed as:

(4.2) $$confidence\ (X \Longrightarrow Y) = \frac{Frequency\ (X \cup Y)}{Frequency\ (X)}$$

In order to draw meaningful conclusions from associations, association rules must meet both a minimum support level (*minsup*) and a minimum confidence level (*minconf*) simultaneously. Low to medium support values often result in a larger number of frequent item sets, making it difficult to identify useful association rules. A measure known as the lift can be used to further filter or rank discovered rules (Brin et al. 1997). Lift is defined as the ratio of the observed support of *X* and *Y* to that expected if *X* and *Y* were independent. Lift can be mathematically represented as

(4.3) $$lift(X \Longrightarrow Y) = \frac{support(X \cup Y)}{support(X)\ support(Y)}$$

With this ratio in mind, lift values greater than 1 imply that observed support of *X* and *Y* is larger than if the two were independent. In terms of association strength, lift values greater than 1 indicate stronger associations (Hahsler et al. 2005).

In order to avoid an exploding number of candidate bundles and to minimize the risk of discovering spurious associations, we apply the Apriori algorithm as suggested by Agrawal and Srikant (1994) and Yabing (2013). Generally, the Apriori algorithm identifies the frequent individual items in the dataset- i.e. individual items that meet both the *minsup* and *minconf*. These are then expanded to larger item sets (bundles) by sequentially adding more items (practices) and only keeping bundles that meet the *minsup* and *minconf*. Any subsets bundles that do not meet the *minsup* and *minconf* are filtered out and those meeting the *minsup* and *minconf* are integrated in the appropriate frequent bundles. The frequent item sets determined by Apriori can then be used to discover association rules based on general trends in the dataset.

Data

This study uses data from the 2018 Oklahoma Beef Management and Marketing Survey conducted by Oklahoma State University Department of Agricultural Economics in collaboration with the Department of Animal Science and the Oklahoma Cooperative Extension Service (IRB Approval No. AG1743). The United States Department of Agriculture's (USDA) National Agricultural Statistics Service (NASS) was contracted for sampling, distributing, data collection, and data entry. The survey targeted cow-calf producers in Oklahoma’s four regions namely Northwest (NW), Northeast (NE), Southwest (SW), and Southeast (SE). Information obtained by the survey includes producer demographics as well as adoption of value-added production and marketing practices. A total of 5000 surveys were sent with 1,495 surveys completed, resulting in a 29.9% response rate at the state level. Of the survey respondents, 48.76% were mail respondents while 51.24% were respondents to the phone follow-up.

Table 4.1 gives a summary of survey responses by region. A noteworthy point in this table is the strong response rate, which is well above 25% in all regions, as well as the similarity in response rate across all 4 regions. Of the 1,495 respondents who completed the survey, 1,210, representing 81% of respondents, reported being a cow-calf producer in the year under consideration. It is these 1,210 respondents on which this study focuses. The survey elicits information regarding producer adoption of various production and marketing practices. This study focuses on 7 production (management) practices namely castration, dehorning plus polled genetics, deworming, feed bunk training, 45-day weaning period prior to marketing, respiratory vaccinations, and use of implants.

In this study, we apply the same analogy used in MBA to a subset of data that includes only cow-calf producers who adopted at least one practice (1,108). This type of analysis searches for associations among practices, and thus usable data are only those with at least one practice recorded.

Results and Discussion

Table 4.2 presents summary statistics of demographic information for producers who adopted at least one of the seven management practices. In regards to age, 44% of producers are 65 years or older, whereas only 3% are below age 35. This age distribution corresponds with others such as Williams et al. (2013), McBride and Mathew (2011) and USDA-NASS (2007), who find the percentage of producers above age 65 to be between 42% and 48% and those below 35 years old to be about 5% at both the state and national levels. Regarding education, over half (60%) of producers had either completed high school or vocational/technical training, whereas those with college or graduate qualification make up close to 40%. Only 2% reported that they had no formal education. In terms of experience in cow-calf production, results show that slightly over three-fifths of producers have been working in the sector for more than 25 years.

Turning to herd size, results reveal herd size of less than 100 to be the most common, accounting for slightly over 60% of producers, with only about 3% reported owning herd sizes of more than 500. This finding is consistent with other similar surveys in Oklahoma (e.g., USDA-NASS 2007; Vestal et al. 2007) that report the majority of producers owning less than 100 head. The contribution of cattle income to household income shows that cattle accounts for less than 40% of income for about 60% of producers, a finding consistent with Vestal et al. (2007) who report 80% of producers indicated that cattle income accounted for less than 40% of total household income. Williams et al. (2013) find a similar pattern with cattle income accounting for less than 40% among 76% of Oklahoma cow-calf producers.

Table 4.3 gives a summary of adoption rates for 7 calf management practices across all cow-calf producers in the dataset, regardless of whether they adopted any practice or not. Results show varying rates of adoption reported by respondents across practices. Deworming of calves, reported by about 88% of producers, is the most adopted practice, followed closely by castration of young bull calves (84%). This result is somewhat different from Williams et al. (2013) who found castration to be the most widely adopted practice by Oklahoma cow-calf producers with

72% of producers reporting using the practice, while deworming, at a 62% adoption rate, was second. The current analysis not only indicates an increase in both castration and deworming adoption rates, but also that deworming adoption increased by a higher percentage than castration, making deworming the most widely adopted practice. Similar to Williams et al. (2013), slightly less than half (47%) of respondents reported dehorning their calves. However, about 70% of those that reported not dehorning their calves used polled genetics to ensure calves had no horns. Thus, managing horns accounts for about 87% of respondents, implying a high percentage of producers are marketing their calves without horns.

Other widely adopted practices include getting calves used to feeding from feed bunks (78%) and 45-day weaning of calves prior to marketing (66%). Adoption rates for the remaining practices are less than 60% with use of implants being the least adopted at 26%. Low adoption rates for implants are reported in other studies (e.g., see Williams et al. 2013; Johnson et al. 2010).

Figure 4.1 shows practice adoption across 4 regions in addition to state averages, but only for producers who reported adopting at least one practice of the seven: castration, horn management, 45-day weaning period, feed bunks, deworming, respiratory vaccinations, and implants. State level adoption rates (Table 4.2) indicate deworming, managing horns, castration, feed bunk, and 45-day weaning to be more widely adopted practices, whereas respiratory vaccination and implants are the least adopted. Similar to the pattern at the state level (Table 4.2), results in Figure 4.1 show that deworming is more widely adopted, followed by horn management with respiratory vaccination and implants being the least adopted.

Past research (Johnson et al. 2010; Benham et al. 2007; Popp, Faminow and Parsch 1999) has shown that there can be significant variation in practice adoption across different geographic regions, even within the same state. Understanding regional differences is important for improved targeting of extension programs. Results in Figure 4.1 show regional differences in terms of adoption across all 7 practices, with the NW having higher adoption rates across all 7 practices, a finding consistent with Williams et al. (2013). Adoption rates in the other regions appear to vary more than in the NW. A look at two basic practices, horn management and castration, across regions indicates that these practices are adopted by 91% of producers in the NW, and are also the highest across all 4 regions. This finding is similar to Williams et al. (2013) who find dehorning (excluding polled genetics) adoption to be highest in the NW and Panhandle. In the NE, horn management (85%) is higher than castration (78%) and a similar pattern is observed in the SE with horn management at 87% and castration at 81%. The pattern in the SW is similar to NW, where adoption of the two practices is almost equal.

Association Rule Results

A total of 448 association rules are discovered initially. The discovered rules are plotted against support, lift, and confidence (Figure 4.2) to help filter out bundles that fail to meet the minimum parameter (support, confidence, and lift) thresholds. One important point from this plot is the fact that most rules that have high confidence and high lift tend to have low support, implying that bundles in this segment of the plot are not widely adopted in the sample. An inspection of this segment revealed that most bundles included implants as consequent of almost all preconditioning practices (antecedent). Bundles in this segment represent just over 10% of the sub-sample. This is expected given the low adoption rate of implants. However, this result also reveals that when implants are used, they are mostly adopted together with all or close to all preconditioning practices. To have meaningful rules, we filtered out bundles with confidence below 0.5, support below 0.3, and lift below 1.2. With these parameter specifications, 27 rules identifying likely bundles are discovered (Table 4.4 A in the appendix). Four of these rules (15%) include deworming, horn management, and castration.

We used the association rule technique to discover possible joint adoption among practices, with their respective probabilities. Figure 4.3 depicts a series of practice adoption likelihoods at different stages of the adoption hierarchy. The is constructed from the over 400 association rules discovered in the data. Past studies (e.g.,Williams et al. 2013; Schumacher, Schroeder and Tonsor 2012) classify horn management, castration and deworming as basic practices most widely adopted at the minimum. Based on their high adoption rates, horn management, deworming, and castration are designated as base practices in this study. With this in mind, Figure 4.3 illustrates the conditional probabilities for adoption of practices contained in the preconditioning bundle. The base practice bundle of castration, horn management and deworming is assumed to be the antecedent. Conditional probabilities for the consequents of feed bunk training, 45-day weaning, and vaccinations individually are 0.81, 0.71, and0.32, respectively. That is, producers who adopt the antecedent base bundle have a significantly higher conditional probability of also adopting feed bunk training or 45-day weaning than vaccinations as a consequent. This difference in probabilities could be because some producers may opt to wean at 30 and not 45-days prior to marketing, while still training calves to feed from bunks. Under such circumstances, producers adopt a shorter (30 days) weaning period but still introduce feed bunks during the weaning period.

If the antecedent is comprised of base practices and 45-day weaning, there is an 85% conditional likelihood a producer will also adopt feed bunks as a consequent. Comparatively, the conditional likelihood that a producer will adopt 45-day weaning as a consequent of an antecedent consisting of base practices and feed bunks is lower at 75%. Another noteworthy point from Figure 4.3 is the difference in conditional probabilities of vaccination following adoption of an antecedent bundle made up of base practices with either 45-day weaning or feed bunks. The conditional probability of vaccination given an antecedent containing base practices plus 45-day weaning (0.68) is greater than if the antecedent contains base practices and feed bunks (0.62). The conditional probability for vaccination increases further to 70% if both 45-day weaning and feed bunk training plus base practices are in the antecedent bundle. This perhaps indicates that producers understand the health impacts associated with weaning and are more likely to vaccinate calves if they adopt both 45-day weaning and feed bunks in addition to the base. It is interesting to also note that the conditional likelihood of 45-day weaning (0.85) is higher if the antecedent contains base practices plus vaccination than if the antecedent only contains base practices. This same pattern is observed with regards to feed bunks, and reinforces the earlier explanation on producers being aware of health risk associated with weaning, in the sense that producers appear to be more likely to implement 45-day weaning and feed bunks conditional on base practices plus vaccination. Taken together, these results suggest that producers are cognizant of the health benefits of vaccination, particularly if calves undergo 45-day weaning and feed bunk training.

While implants are not necessarily part of preconditioning practices and coupled with the fact that they are less widely adopted than the other practices considered here (26%), we examined the likelihood of implants under three different scenarios: 1) castration only as an antecedent; 2) three base practices as antecedent; and 3) all 6 preconditioning practices as antecedent. Results of this analysis (Figure 4.4) indicate the conditional probability of implant adoption increasing with the number of practices already adopted. With the basic 3 as antecedent (already adopted), the probability of adopting implants is 0.32, an increase of 5 percentage points from the scenario where castration only is the antecedent. When all 6 preconditioning practices are being implemented, the conditional probability of implant adoption further increases to 0.42. One implication of this result is that efforts to improve adoption of implants may benefit from initially targeting producers who already adopted preconditioning practices or are close to doing so.

Conclusion

This study set out to identify practices likely to be jointly adopted among cow-calf management practices recommended for value addition to calves and to measure the conditional probability of adoption of specific practice bundles. In terms of individual practice adoption, results show varying adoption rates across all seven practices, with deworming (87.5%) being the most widely adopted, followed by horn management (86.5%) and castration (83.5%). The least adopted practice is use of implants adopted only by 26% of producers in the sample. Overall, results also reveal variations in practice adoption rates across geographical regions, with the Northwest having higher adoption rates across most practices.

Using a sub-sample of producers who reported adopting at least one practice and applying the association rule learning technique, and in particular the market basket analysis with Apriori algorithm, we find varying conditional probabilities for joint adoption among practices. In particular, results show that horn management, deworming, and castration are the more adopted practices and also tend to be jointly adopted as a bundle in combination with other practices. Using horn management, deworming, and castration as the base, we construct a hierarchy of practice adoption leading to preconditioning. Results show that the probability of adding 45-day weaning to an antecedent bundle containing base practices and feed bunks is lower than that of adding feed bunks to an antecedent bundle comprised of base practices and 45-day weaning. One plausible explanation for this result is that producers might choose to wean calves for a shorter period, e.g. 30 days, yet still adopt feed bunk training during the weaning period. We also find likelihood of vaccination to be higher if a producer adopts both feed bunks and 45-day weaning in addition to base practices. Further, base practices in combination with vaccination increase the conditional likelihood of 45-day weaning and feed bunks, further supporting the explanation that producers recognize the importance of vaccination in helping calves go through the 45-day weaning and feed bunk training. Thus efforts to help improve adoption of vaccination may consider targeting producers that are more likely to implement base practices and both 45-day weaning and feed bunks, while also designing programs to encourage adoption of these antecedent practices.

With regards to implants, we evaluated the likelihood of adoption under three different scenarios; with castration only; with horn management, deworming, and castration; and with the six preconditioning bundle practices. Results show low implant adoption likelihood when only castration is the antecedent. The likelihood increases when deworming and horn management are added to the antecedent and is highest when the antecedent has all preconditioning practices. This perhaps reflects the low value producers attach to implants and thus will only adopt when they are already implementing sll other practices. It may also reflect a misunderstanding of the value of non-implanted cattle at marketing if those animals are not in a process verified program. One implication of this result is that extension efforts aimed at improving implant adoption may be more successful by focusing on producers who have adopted all preconditioning practices or are close to doing so. At the same time, programs encouraging adoption of all or most preconditioning practices should target producers lagging in terms of achieving preconditioning.

This study attempted to identify practice bundles most often adopted among producers, with the aim of helping improve extension targeting as well as practice adoption decision modeling. Existing studies on adoption decisions assume independence among practices, although this may not always hold true. Having identified possible hierarchies and likely practice bundles, future studies on adoption decision modeling might benefit from considering the constructed bundles as a starting point, thus relaxing the independence assumption. Model estimates from decision models with and without practice independence assumption can then be compared.

References

Agrawal, R., H. Road, and S. Jose. 1993. "Mining Association Rules between Sets of Items in Large Databases." In *ACM SIGMOD Conference Washington DC. USA*. pp. 1–10.

Agrawal, R., and R. Srikant. 1994. "Fast Algorithms for Mining Association Rules." In *Proceedings of The 20th International Conference on Very Large Data Bases*, pp. 487-499.

Avent, R.K., C.E. Ward, and D. Lalman. 2004. "AGEC-583 Economic Value of Preconditioning Feeder Calves." Oklahoma State University. Available at http://osuextra.com/pdfs/F-583web.pdf, accessed March 2019.

Behrends, L., T.G. Field, and K. Conway. 2001. "The Value of Information as Perceived by Feedlot Managers." *Animal Sciences Research Report, Colorado State University*.

Benham, B.L., A. Bracia, S. Mostaghimi, J.B. Lowery, and P.W. McClellan. 2007. "Comparison of Best Management Practice Adoption Between Virginia's Chesapeake Bay Basin and Southern Rivers Watersheds." *Journal of Extension*, 45(2).

Brin, S., R. Motwani, J.D. Ullman, and S. Tsur. 1997. "Dynamic Itemset Counting and Implication Rules for Market Basket Data." *ACM SIGMOD Record* 26(2):255–264.

Chen, Y.-L., K. Tang, R.-J. Shen, and Y.-H. Hu. 2005. "Context-Based Market Basket Analysis in a Multiple-Store Environment." *Decision Support Systems* 45(1):150–163.

Dhuyvetter, K.C., A.M. Bryant, D.A. Blasi, N.C. Edwards, M.R. Langemeier, J.B. Morgan, and C.W. Shearhart. 2005. "Preconditioning Beef Calves: Are Expected Premiums Sufficient to Justify the Practice?" *The Professional Animal Scientist* 21(6):502–514.

Hahsler, M., B. Grün, and K. Hornik. 2005. "arules - A Computation Environment for Mining Association Rules and Frequent Item Sets." *Journal of Statistical Software* 14(15):1–25.

Hipp, J., U. Güntzer, and G. Nakhaeizadeh. 2007. "Mining Association Rules: Deriving a Superior Algorithm by Analyzing Today's Approaches." In *European Conference on*

Principles of Data Mining and Knowledge Discovery (pp. 159-168). Springer, Berlin, Heidelberg.

Johnson, R.J., D. Doye, D.L. Lalman, D.S. Peel, K. Curry Raper, and C. Chung. 2010. "Factors Affecting Adoption of Recommended Management Practices in Stocker Cattle Production." *Journal of Agricultural and Applied Economics* 42(1):15–30.

Lalman, D., and G. Mourer. 2014. "Effects of Preconditionning on Health, Performances and Prices of Weaned Calves." Extension Fact Sheet F-3529, Oklahoma State University.

Linoff, G.S., and J.A. Berry, Michael. 2011. *Data Mining Techniques for Marketing, Sales, and Customer Relationship Management* 3rd ed. Hoboken, NJ: Wiley Computer Publishing.

McBride, W.D., and K.J. Mathew. 2011. "Diverse Structure and Organization of US Beef Cow-Calf Farms." No. 102764. United States Department of Agriculture, Economic Research Service.

Popp, M.P., M.D. Faminow, and L.D. Parsch. 1999. "Factors Affecting the Adoption of Value-added Production on Cow-Calf Farms." *Journal of Agricultural and Applied Economics* 31(01):97–108.

Popp, M.P., M.D. Faminow, and L.D. Parsch. 1999. "Factors Affecting the Adoption of Value-Added Production on Cow-Calf Farms." *Journal of Agricultural and Applied Economics* 31(September 2016):97–108.

Pruitt, J.R., J.M. Gillespie, R.F. Nehring, and B. Qushim. 2012. "Adoption of Technology, Management Practices, and Production Systems by U.S. Beef Cow-Calf Producers." *Journal of Agricultural and Applied Economics* 2(May):203–222.

Schumacher, S., K.C. Raper, and D.S. Peel. 2017. "Demographic Influences on Nonadoption of Calf Management and Marketing Practices for Cow-Calf Operations." *Journal of Applied Farm Economics* 1(2).

Schumacher, T., T.C. Schroeder, and G.T. Tonsor. 2012. "Willingness-to-Pay for Calf Health Programs and Certification Agents." *Journal of Agricultural and Applied Economics*

44(2):191–202.

Shaw, M.J., C. Subramaniam, G.W. Tan, and M.E. Welge. 2001. "Knowledge Management and Data Mining for Marketing." *Decision Support Systems* 31(1):127–137.

United States Department of Agriculture - National Agriculture Statistics Services (USDA-NASS). 2009. "2007 Census of Agriculture: United States Summary and State Data." Available at: http://www.agcensus.usda.gov/Publications/2007/Full Report/usv1.pdf

Vestal, M., C. Ward, D. Doye, and D. Lalman. 2007. "Cow-Calf Production Practices in Oklahoma – Part 1." Oklahoma Cooperative Extension Service. AGEC-245.

Ward, C., M. Vestal, D. Doye, and D.L. Lalman. 2008. "Factors Affecting Adoption of Cow-Calf Production Practices in Oklahoma." *Journal of Agricultural and Applied Economics* 40(3):851–863.

Ward, H., J. Powell, and S. Gadberry. 2019. "Preconditioning Programs for Beef Calves." No. FSA3074, Cooperative Extension Service, University of Arkansas, Division of Agriculture.

Williams, B.R., K.C. Raper, E. A. DeVuyst, D. Peel, D. Lalman, and C. Richards. 2013. "Demographic Factors Affecting the Adoption of Multiple Value-Added Practices by Oklahoma Cow-Calf Producers." *Journal of Extension* 51(6).

Williams, G.S., K.C. Raper, E.A. DeVuyst, D. Peel, and D. McKinney. 2012. "Determinants of Price Differentials in Oklahoma Value-Added Feeder Cattle Auctions." *Journal of Agricultural and Resource Economics* 37(1):114–127.

Yabing, J. 2013. "Research of an Improved Apriori Algorithm in Data Mining Association Rules." *International Journal of Computer and Communication Engineering* 2(1):25–27.

Table 4.1. Survey Response Summary

Quadrant	**Sampled**	**Returned**	**Percent Response Rate**
NW	1115	331	29.7
NE	1655	531	32.1
SE	980	265	27.0
SW	1250	368	29.4
State	5000	1495	29.9

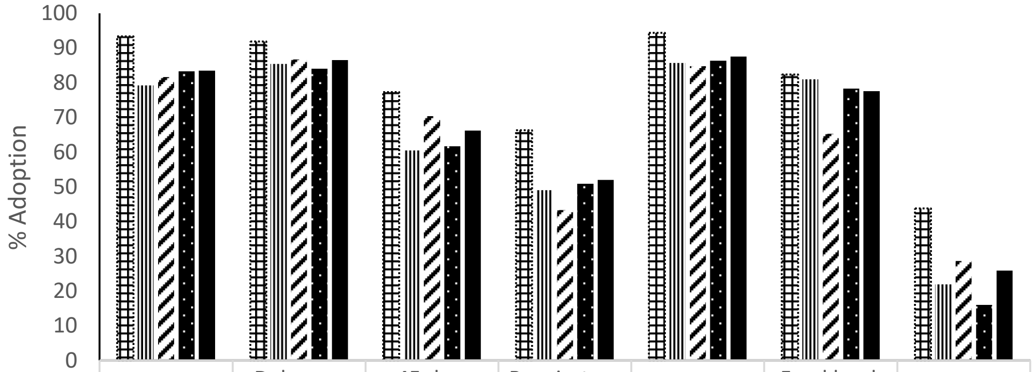
Table 4.2. Summary Statistics for Producer Demographics

Variable	Number of Producers	Percent
Age		
<=25	2	0.2
25-29	9	0.9
30-34	19	1.89
35-39	34	3.39
40-44	45	4.49
45-49	59	5.88
50-54	86	8.57
55-64	288	28.71
65-74	283	28.22
>= 75	158	15.75
N/A	20	1.99
Education Level		
High school	392	39.08
Vocational/technical	198	19.74
Bachelor’s degree	219	21.83
Graduate degree	150	14.96
None	23	2.29
N/A	21	2.09
Years in Cattle Production		
<=5	36	3.59
5 - 10	75	7.48
11 - 15	86	8.57
16 - 20	78	7.78
21 - 25	88	8.77
>=25	624	62.21
N/A	16	1.6
Herd Size		
1 - 24	160	15.95
25 - 49	195	19.44
50 - 99	244	24.33
100 - 249	270	26.92
250 - 499	88	8.77
500 - 749	26	2.59
750 - 999	5	0.5
1000+	8	0.8
N/A	7	0.7
Total Income		

Variable	Number of Producers	Percent
<=$30,000	77	7.68
$30K - $59,999	195	19.44
$60K - $89,999	254	25.32
$90K - $119,999	148	14.76
>=120K	220	21.93
N/A	109	10.87
Percent Cattle Income		
0	66	6.58
1 - 20	375	37.39
21 - 40	201	20.04
41 - 60	133	13.26
61 - 80	86	8.57
81 - 100	75	7.48
N/A	67	6.68
BQA Training		
Yes	98	9.77
No	871	86.84
N/A	34	3.39
OSU Master Cattleman Program		
Yes	45	4.49
No	910	90.73
N/A	48	4.79
Local/Count Cattleman member		
Yes	214	21.34
No	789	78.66
National Cattleman Beef Association		
Yes	69	6.88
No	934	93.12

Table 4.3. Practice Adoption Rates

Practice	Number of Respondents	Number of Respondents Reporting Adoption	Adoption Rate (%)
Castration	947	791	83.53
Polled genetics	450	317	70.44
Dehorn + polled genetics	990	857	86.57
45-day weaning	890	589	66.18
Respiratory vaccines	889	462	51.97
Deworm	923	808	87.54
Feed bunk training	898	697	77.62
Implant	860	223	25.93



	Castration	Dehorn + polled	45 day weaning	Respiratory vaccines	Deworm	Feed bunk training	Implant
NW	93.39	91.93	77.30	66.30	94.39	82.42	43.79
NE	79.24	85.40	60.50	49.04	85.71	80.98	21.94
SE	81.61	86.67	70.35	43.35	84.80	65.32	28.65
SW	83.26	84.03	61.68	50.92	86.36	78.34	16.04
State	83.53	86.57	66.18	51.97	87.54	77.62	25.93

NW NE SE SW State

Figure 4.1. Practice Adoption among Adopters by Region

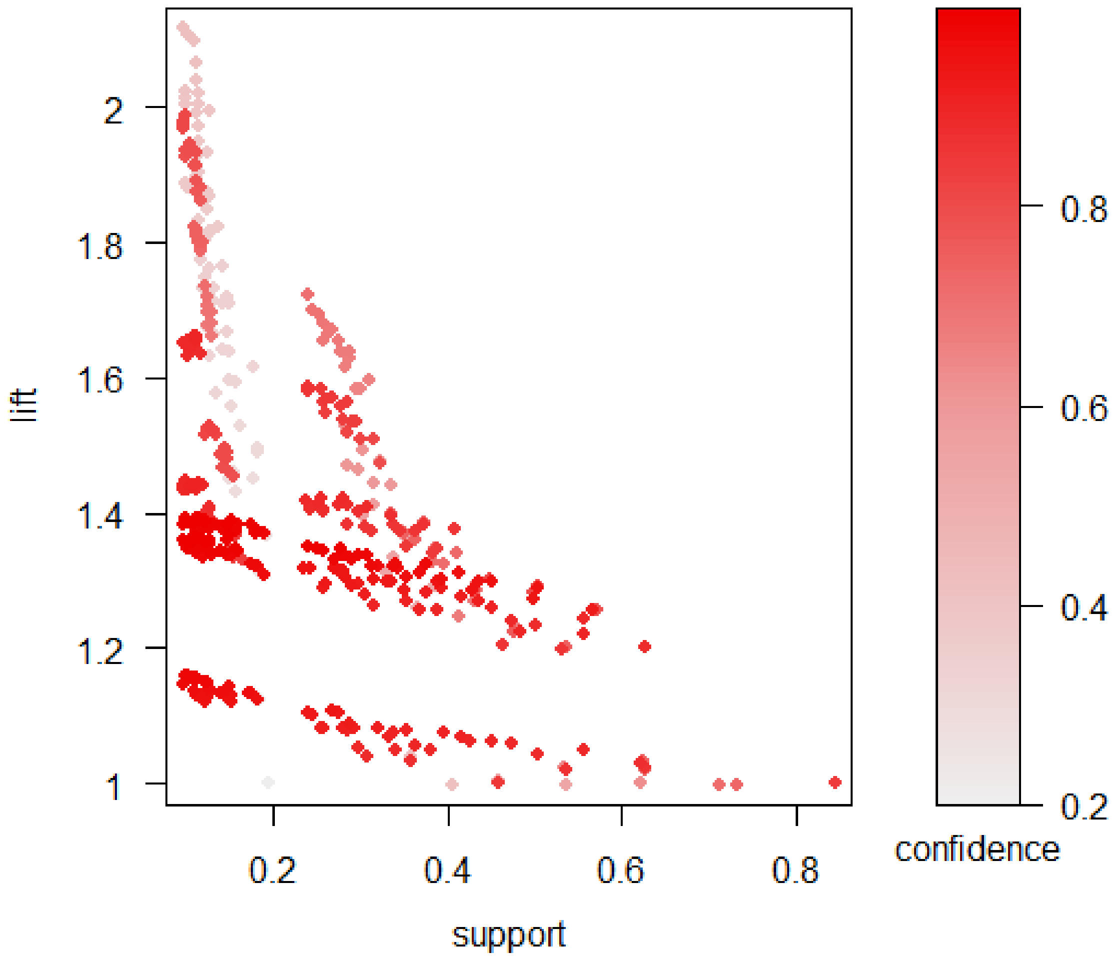


Figure 4.2. Scatter Plot of Association Rules by Support, Confidence and Lift

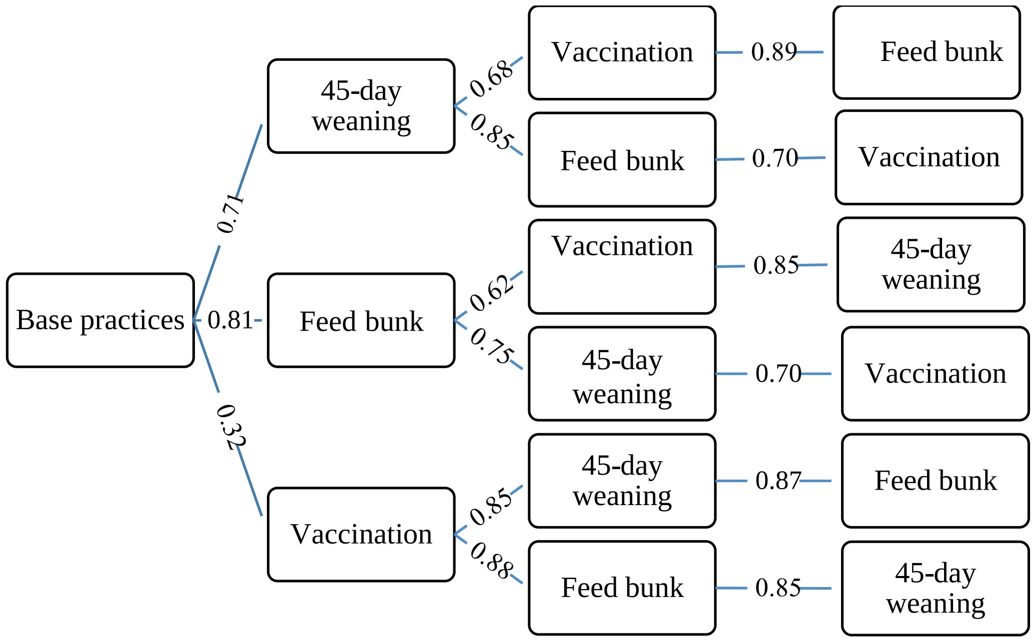


Figure 4.3. Probability of Joint Adoption for Selected Antecedent and Consequent Management Practice Combinations[7]

[7] Base practices are deworming, horn management, and castration

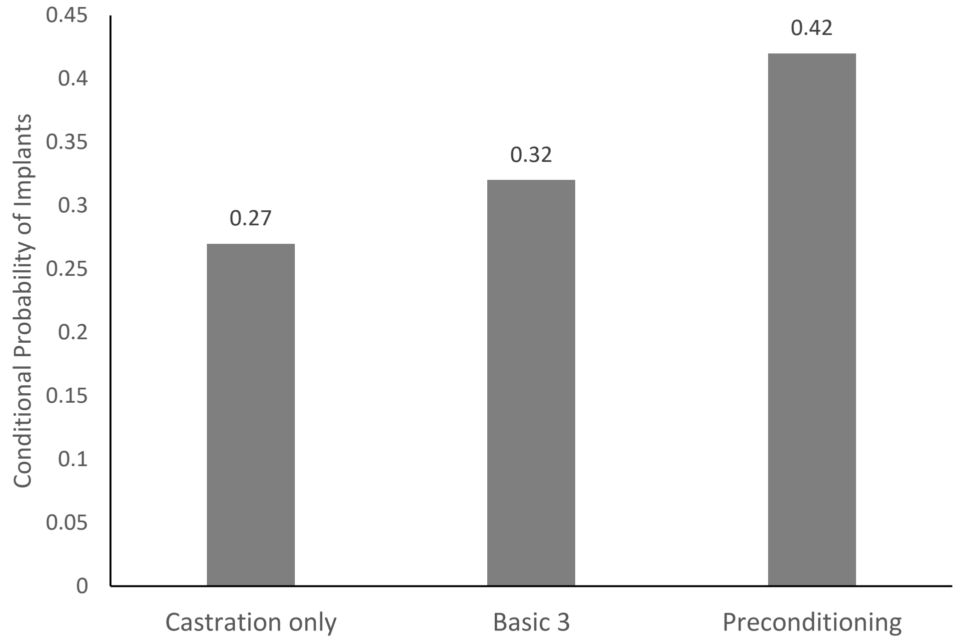


Figure 4.4. Probability of Implant Adoption under Three Antecedent Scenarios

.

CHAPTER V

CONCLUSIONS

The three essays presented in this dissertation are all concerned with producer decision making. The first essay focuses on input application decision under yield uncertainty due to WSM infection; the second essay constructs a feedlot cost of gain prediction model to help improve ex-ante cost of gain prediction and thus breakeven price estimation for varying sizes of feeders; and the third sheds light on cow-calf value enhancement practices likely to be jointly adopted.

In the first essay, results show exponential yield decline with increasing WSM infection severity – as measured by tissue reflectance values- consistent with similar studies. However, this study represents the first analysis to estimate the economic threshold of the WSM during the growing season. Estimates indicate the threshold reflectance to range from about 8.3 to 9.5, for readings taken around late April to early May, and are sensitive to input and output prices used in the construction of partial budgets. With about 14 percent of the sampled plots recording reflectance values greater than the threshold, it can be inferred that producers whose crop is infected with WSM may potentially save resources and salvage some value by discontinuing input applications to the crop and grazing out the wheat forage from infected fields whose infection severity exceeds the estimated threshold values at a particular time.

However, prioritizing good management practices such as ridding the field of volunteer wheat and weeds early enough before planting, to eliminate the green bridge, still remains critical in helping reduce income losses.

In the second essay, results show that the constructed COG prediction model outperforms the naïve model in terms of prediction accuracy. Unlike the naïve model which assumes future closeout COG to be a function of current closeout COG only, the constructed model indicate that in addition to current closeout COG, expected corn price and its lagged values, as well as seasonality are important predictors of closeout COG. In terms of seasonality, results suggests that earlier placements in the year (winter and spring) tend to be associated with lower COG relative to later placements. Although the constructed model predicts COG better, breakeven price calculated using the model's prediction does not significantly differ from that calculated using naïve model predictions. Results show a reduction in spreads between 800-850 pound feeder market price and breakeven price for lighter weights mostly in the first half of the year. This result perhaps indicates that COG does not completely explain the difference between feeder market price and feedlot breakeven price, because if it did, then the spread based on the constructed model estimates would have been significantly reduced. It is possible that other market conditions not captured by COG, such as changes in feeder supply and demand overtime do explain some of the price spread. Although the constructed model may not capture broader market conditions, it does account for expected changes in corn demand and supply as reflected in the corn futures price. Given the significance of corn in feedlot rations, it is reasonable to expect the COG prediction from the constructed model to be more dynamic with improved accuracy.

Taken together, results of this study show that COG is not the major component explaining spreads between 800-850 pound steer feeders market price and feedlot breakeven price for lighter weight steers, when corn prices fluctuate only marginally. Much of the spread is most likely explained by the feeder cattle supply and demand dynamics across time. Thus one important caveat of our model is the inability to capture changes in supply and demand of feeder cattle across time. This is because our model relies on the strong relationship between corn price and COG, and changes in feeding efficiency across seasons. Where feeder supply and demand data are available, future research should consider extending our model by accounting for such changes across weight and time in the calculation of feedlot breakeven price.

The third essay examined the joint adoption of recommended cow-calf management practices among Oklahoma producers. Results show that the most widely adopted practices are deworming, horn management, and castration each adopted by more than 80% of respondents. Implants are the least adopted only accounting for 26% of respondents. Using a sub-sample of producers who reported adopting at least one practice and applying the association rule learning technique, in particular the market basket analysis with Apriori algorithm, results show varying likelihood of joint adoption among practices. With horn management, deworming, and castration as the base, we construct a hierarchy of practice adoption likely leading to preconditioning. The probability of adding 45-day weaning to an antecedent bundle containing base practices and feed bunks is lower than that of adding feed bunks to an antecedent bundle comprising base practices and 45-day weaning. The possibility of producer choosing to wean for a shorter period e.g., 30 days, yet still adopt feed bunk training during weaning period could explain this outcome.

Results also show a higher likelihood of adopting vaccination given an antecedent comprising base practices and both 45-day weaning and feed bunks. In a reciprocative manner, vaccination, if added to the base practice, increases conditional likelihood of both 45-day weaning and feed bunk. As for implants, results indicate that likelihood of adoption is highest when the antecedent contains all preconditioning practices. The implication for extension is that efforts aimed at improving implant adoption may be more successful by focusing on producers who have adopted all preconditioning practices or are close to doing so. At the same time, programs encouraging adoption of all or most of preconditioning practices should be targeted at producers lagging in terms of achieving preconditioning.

As this study has identified possible hierarchies and likely practice bundles, future studies should model adoption decision using constructed bundles, thus relaxing the independence assumption. Model estimates from decision models with and without practice independence assumption can then be compared.

APPENDICES

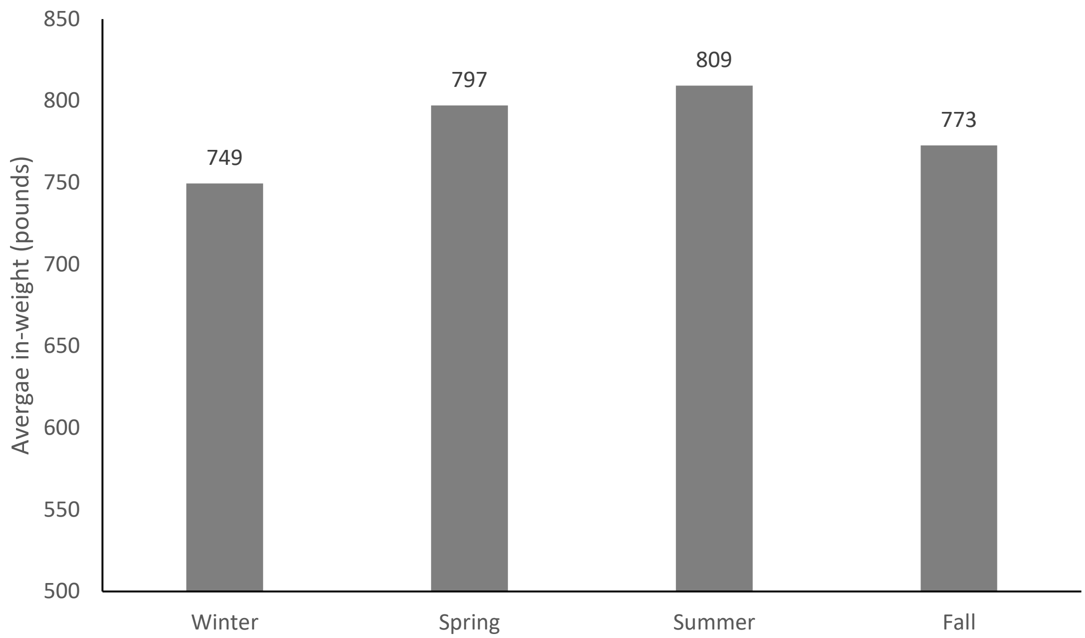


Figure 3.1A. Average In-weight by Placement Season

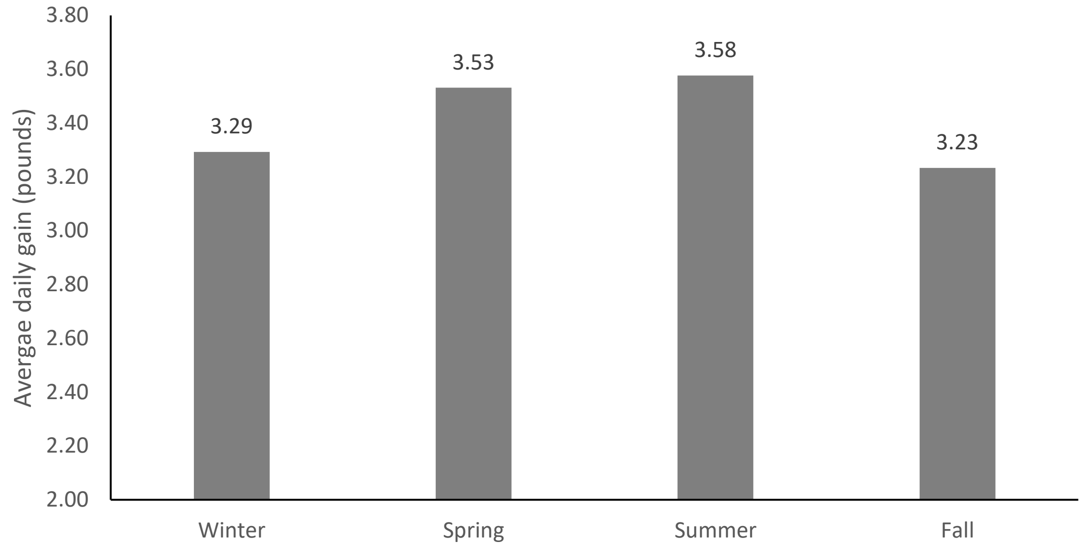


Figure 3.2A. Average Daily Gain by Placement Season

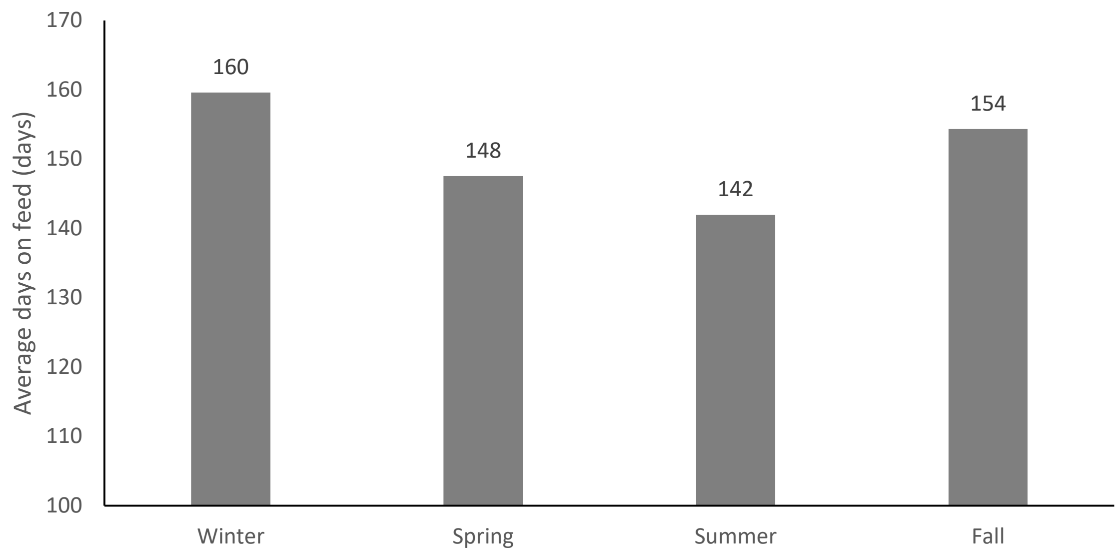


Figure 3.3A. Average Days on Feed by Placement Season

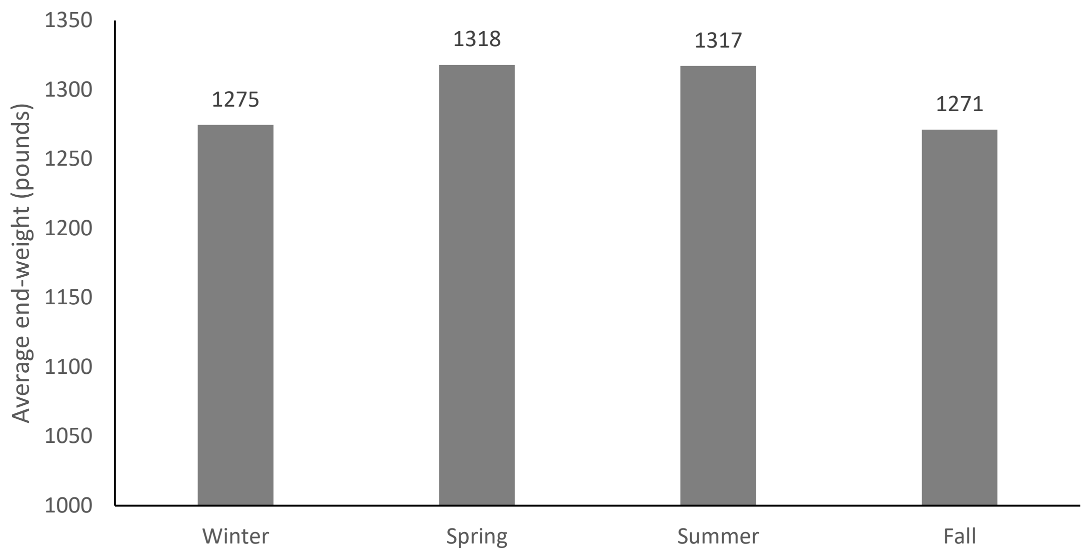


Figure 3.4A. Average End-Weight by Placement Season

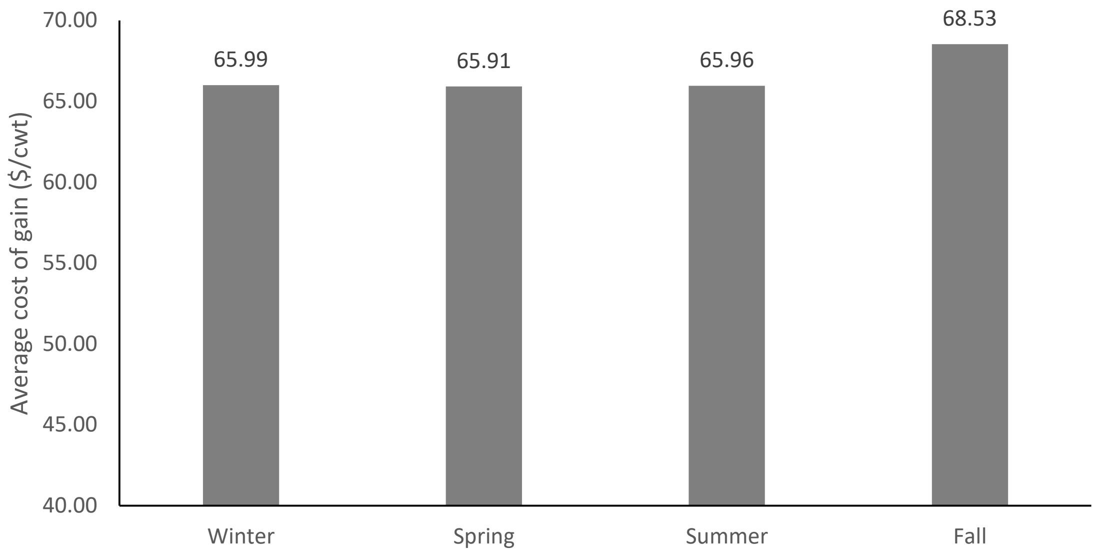


Figure 3.5A. Average Cost of Gain by Placement Season

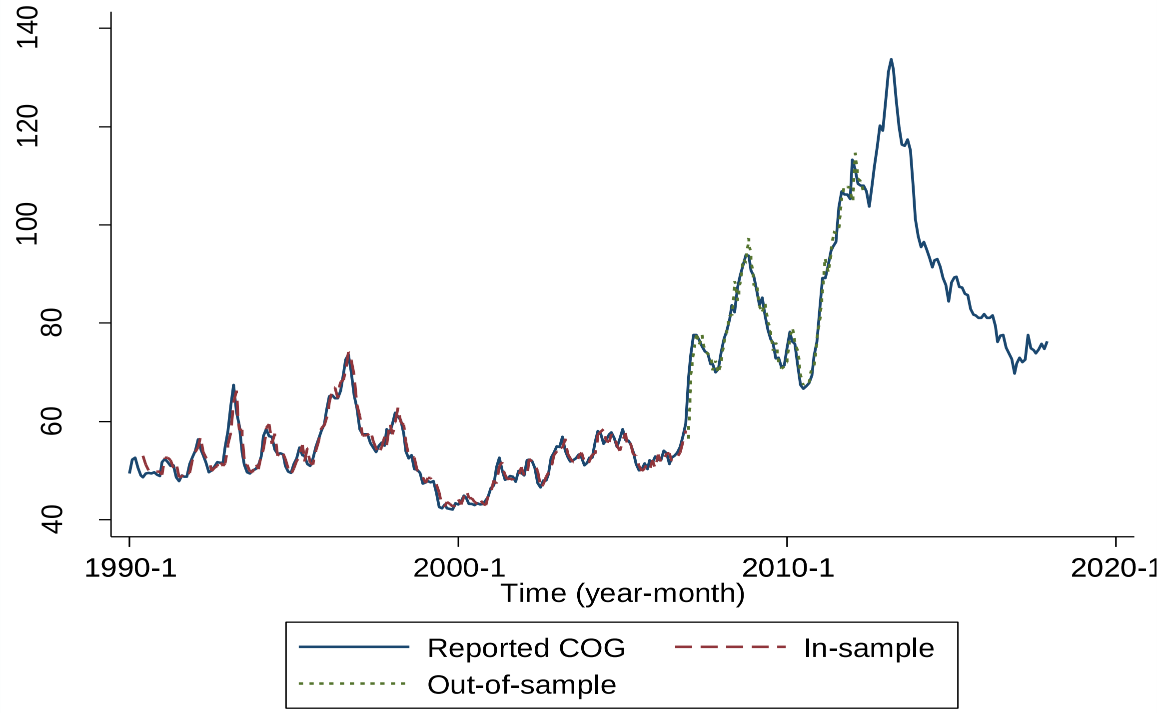


Figure 3.6A. Out-of-Sample Prediction of Cost of Gain, January 2006 – June 2012

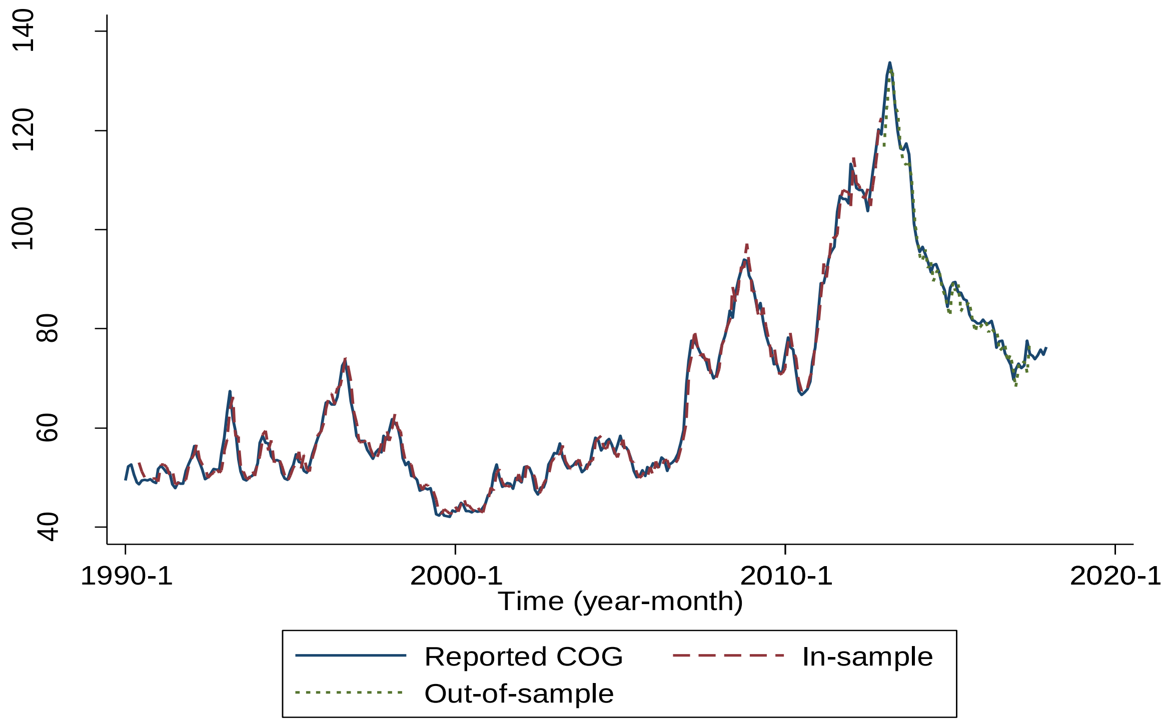


Figure 3.7A. Out-of-Sample Prediction of Cost of Gain, January 2012-June 2017

Table 4.1 A. Practice Adoption Association Rules

Antecedent		Consequent	Support	Confidence	Lift	Count
{feed bunk,castration,horn management ,vaccination}	=>	{deworm}	0.28	0.98	1.34	310
{feed bunk,castration,vaccination}	=>	{deworm}	0.31	0.97	1.33	340
{castration,horn management ,vaccination,45-day weaning }	=>	{deworm}	0.27	0.97	1.33	299
{castration,horn management ,vaccination}	=>	{deworm}	0.32	0.96	1.32	353
{castration,horn management ,45-day weaning }	=>	{deworm}	0.40	0.95	1.30	438
{horn management ,deworm,vaccination,45-day weaning }	=>	{castration}	0.27	0.94	1.32	299
{horn management ,vaccination,45-day weaning }	=>	{castration}	0.28	0.94	1.32	307
{feed bunk,horn management ,deworm,vaccination}	=>	{castration}	0.28	0.94	1.31	310
{horn management ,deworm,vaccination}	=>	{castration}	0.32	0.93	1.30	353
{feed bunk,horn management ,deworm,45-day weaning }	=>	{castration}	0.34	0.93	1.30	376
{horn management ,deworm,45-day weaning }	=>	{castration}	0.40	0.92	1.29	438
{castration,horn management ,deworm,vaccination}	=>	{feed bunk}	0.28	0.88	1.41	310
{castration,horn management ,deworm,45-day weaning }	=>	{feed bunk}	0.34	0.86	1.38	376
{castration,horn management ,deworm,vaccination}	=>	{45-day weaning }	0.27	0.85	1.57	299
{castration,horn management ,45-day weaning }	=>	{feed bunk}	0.35	0.84	1.35	388
{horn management ,deworm,vaccination}	=>	{45-day weaning }	0.29	0.84	1.56	318
{feed bunk,castration,deworm,vaccination}	=>	{45-day weaning }	0.26	0.84	1.56	285
{castration,horn management ,vaccination}	=>	{45-day weaning }	0.28	0.84	1.56	307
{feed bunk,castration,vaccination}	=>	{45-day weaning }	0.26	0.83	1.54	290
{feed bunk,deworm,vaccination}	=>	{45-day weaning }	0.28	0.83	1.54	308
{horn management ,vaccination}	=>	{45-day weaning }	0.30	0.83	1.54	327
{castration,deworm,vaccination}	=>	{45-day weaning }	0.29	0.83	1.54	321
{castration,horn management ,deworm}	=>	{feed bunk}	0.45	0.81	1.30	499
{horn management ,deworm}	=>	{feed bunk}	0.50	0.80	1.28	554
{feed bunk,castration,horn management ,deworm}	=>	{45-day weaning }	0.34	0.75	1.40	376
{castration,horn management ,deworm}	=>	{45-day weaning }	0.40	0.71	1.32	438
{castration,horn management ,deworm}	=>	{vaccination}	0.32	0.57	1.41	353
{castration,horn management ,deworm, vaccination, 45-day weaning}	=>	{feed bunk}	0.23	0.87	1.42	265
{castration,horn management ,deworm,feed bunk,vaccination}	=>	{45-day weaning}	0.24	0.85	1.59	265

Oklahoma State University Institutional Review Board

Date:	Tuesday, August 8, 2017
IRB Application No	AG1743
Proposal Title:	2017 Oklahoma Beef Management and Marketing Survey
Reviewed and Processed as:	Exempt

Status Recommended by Reviewer(s): Approved Protocol Expires: 8/7/2020

Principal Investigator(s):

Kellie Raper
514 Ag Hall
Stillwater, OK 74078

Derrell Peel
Stillwater, OK 74078

The IRB application referenced above has been approved. It is the judgment of the reviewers that the rights and welfare of individuals who may be asked to participate in this study will be respected, and that the research will be conducted in a manner consistent with the IRB requirements as outlined in section 45 CFR 46.

☐ The final versions of any printed recruitment, consent and assent documents bearing the IRB approval stamp are attached to this letter. These are the versions that must be used during the study.

As Principal Investigator, it is your responsibility to do the following:

1. Conduct this study exactly as it has been approved. Any modifications to the research protocol must be submitted with the appropriate signatures for IRB approval. Protocol modifications requiring approval may include changes to the title, PI advisor, funding status or sponsor, subject population composition or size, recruitment, inclusion/exclusion criteria, research site, research procedures and consent/assent process or forms.
2. Submit a request for continuation if the study extends beyond the approval period. This continuation must receive IRB review and approval before the research can continue.
3. Report any adverse events to the IRB Chair promptly. Adverse events are those which are unanticipated and impact the subjects during the course of the research; and
4. Notify the IRB office in writing when your research project is complete.

Please note that approved protocols are subject to monitoring by the IRB and that the IRB office has the authority to inspect research records associated with this protocol at any time. If you have questions about the IRB procedures or need any assistance from the Board, please contact Dawnett Watkins 219 Scott Hall (phone: 405-744-5700, dawnett.watkins@okstate.edu).

Sincerely,

Hugh Crethar, Chair
Institutional Review Board

VITA

Brian P. Mulenga

Candidate for the Degree of

Doctor of Philosophy

Thesis: ECONOMIC THRESHOLDS OF WHEAT STREAK MOSAIC, FEEDLOT COST OF GAIN PREDICTION, AND JOINT ADOPTION OF COW-CALF PRODUCTION PRACTICES

Major Field: Agricultural Economics

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Agricultural Economics at Oklahoma State University, Stillwater, Oklahoma in May, 2019

Completed the requirements for the Master of Science in Agricultural and Applied Economics at University of Malawi, Lilongwe, March, 2011

Completed the requirements for the Bachelor of Agriculture Science in Agricultural Economics at the University of Zambia, Lusaka, in 2007

Experience:

Graduate Research Assistant in the Department of Agricultural Economics at Oklahoma State University, Stillwater, Oklahoma, August 2016 - May 2019

Research Associate, Indaba Agricultural Policy Research Institute, Lusaka, Zambia, January 2012 – September 2015

In-country coordinator, Regional Hunger and Vulnerability Program, Lusaka, Zambia, September 2007 - August 2008

Professional Memberships:
Economics Association of Zambia
Agricultural and Applied Economics Association
Association of Environmental and Resource Economists