2023

THREE ESSAYS ON THE U.S. BEEF SUPPLY CHAIN: PRODUCTION, MARKETING, AND PRICE DYNAMICS

Erdal Erol
University of Kentucky, erdalerol.e@gmail.com
Author ORCID Identifier:
https://orcid.org/0000-0002-8946-4295
Digital Object Identifier: https://doi.org/10.13023/etd.2023.106

Erdal Erol, Student

Dr. Carl R. Dillon, Major Professor

Dr. Tyler B. Mark, Director of Graduate Studies

THREE ESSAYS ON THE U.S. BEEF SUPPLY CHAIN: PRODUCTION, MARKETING, AND PRICE DYNAMICS

DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Agriculture, Food and Environment
at the University of Kentucky

By

Erdal Erol

Lexington, Kentucky

Director: Dr. Carl R. Dillon, Professor of Agricultural Economics

Lexington, Kentucky

2023

https://orcid.org/0000-0002-8946-4295

ABSTRACT OF DISSERTATION

THREE ESSAYS ON THE U.S. BEEF SUPPLY CHAIN:
PRODUCTION, MARKETING, AND PRICE DYNAMICS

This dissertation contains three essays on select economic components of the U.S. beef industry. The first and second essays concentrate on the different economic problems in beef cattle production. The third essay evaluates the price dynamics and the impact of COVID-19 along the beef supply chain.

The first essay explores the economics of culling decisions in cow-calf operations in the U.S. with a novel application of a dynamic mathematical programming model. The results provide an optimal culling strategy under the base model and a range of optimal strategies that vary with respect to different components such as fertility probabilities, market prices, production and replacement heifer costs, calf weights, and pregnancy check use. The results suggest that producers should cull all cows that are older than age 10 considering their productivity and production costs in light of base product prices. The model recommends culling open cows earlier (at age 7) given their productivity status and probabilities. To measure the sensitivity of the optimal results with respect to components, several experiments are run, and outcomes underline the sensitivity of the optimal strategies to market conditions, cost structure, cow fertility, and pregnancy check use.

The second essay aims to contribute to the U.S. beef cattle price forecasting literature with its model selection framework which compares traditional time series techniques and machine learning algorithms to select the best technique to provide one-week-ahead steer, heifer, and cull cow price forecasts. The study performs these techniques using weekly Kentucky cattle auction prices with lagged variables and dummy variables for weekly seasonal structure. The results demonstrate that while ARIMA models without seasonality has better performance in forecasting steer prices, the LASSO regression provides better forecasts for heifer and cull cow prices. The model selection results point to the superiority of machine learning techniques over standard ARIMA models when forecasting U.S. livestock prices in larger samples.

The third essay investigates the price dynamics along the U.S. beef supply chain and the impact of the COVID-19 shock on the dynamics of vertical price transmission using monthly farm, wholesale, and retail prices for the period 1970-2021. A vertical error correction model along with historical decomposition graphs is employed to measure the impact of the pandemic on price adjustment. The results reveal that the impact of COVID-

19 has been uneven across the beef marketing channel, with farmers taking the burden of the shock. The results underline that in the case of the COVID-19 shock, wholesale prices adjusted more quickly than both farm (threefold) and retail prices (tenfold). Historical decomposition graphs also show that the COVID-19 pandemic caused retailers and wholesalers to have higher prices, while farmers received lower prices than their predicted values. The results indicate that the U.S. beef markets were resilient enough to absorb the shocks and return to their pre-shock patterns in 4 to 6 months.

KEYWORDS: Beef-Cattle Supply Chain, Culling Decisions, Forecasting, Machine Learning, Price Transmission, COVID-19

Erdal Erol

April 21, 2023

Date

THREE ESSAYS ON THE U.S. BEEF SUPPLY CHAIN:
PRODUCTION, MARKETING, AND PRICE DYNAMICS

By
Erdal Erol

Dr. Carl R. Dillon
Director of Dissertation

Dr. Tyler B. Mark
Director of Graduate Studies

April 21, 2023
Date

ACKNOWLEDGMENTS

I would like to extend my deepest gratitude to all those who have contributed to the successful completion of this dissertation.

First and foremost, I am grateful to my advisor, Dr. Carl Dillon, for his invaluable guidance and encouragement during my PhD program. His expertise, feedback, and support have contributed remarkably to both my research and career.

I am also indebted to my committee members, Dr. Kenny Burdine, Dr. Sayed Saghaian, and Dr. Nicholas Pates. Their feedback, critiques, and discussions helped me to shape my research and ensure the quality of my dissertation. I am also grateful to my committee member, Dr. Les Anderson, whose insight and perspective on the beef cattle industry contributed significantly to my dissertation. I would also like to thank my committee member, Dr. Felipe Benguria for providing valuable feedback on my dissertation.

Finally, I am grateful to my family and friends for their unwavering support and trust in me. Their faith in me has been a constant source of motivation throughout my academic journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

1.1 Overview of the U.S. Beef Supply Chain

The United States is the largest producer and consumer of beef in the world, accounting for 20% of the global beef production (USDA Foreign Agricultural Service, 2022). In 2020, total beef production in the U.S. reached 27.2 billion pounds and the number of beef cows was 31.3 million (Tables 1.1 and 1.2). While the U.S. beef-cattle industry is geographically diversified, the top ten states with the highest cattle inventory in 2020 were Texas, Oklahoma, Missouri, Nebraska, South Dakota, Kansas, Montana, Kentucky, North Dakota, and Arkansas (Table 1.3). Although the U.S. has one of the largest cow herds and the largest beef production amounts in the world, it was a net importer in the beef trade with an export amount of 3.0 million pounds and an import amount of 3.3 million pounds in 2020 (USDA Economic Research Service, 2022). Figures 1.1 and 1.2 display trends in the beef and cattle trade for the period 2000-2020. The U.S. beef trade had trade deficits for most of the period with the highest amounts observed between 2004-2007 due to the impact of Bovine Spongiform Encephalopathy (BSE) disease in the U.S. The United States was a net importer of live cattle during the period.

The U.S. beef supply chain is a complex and dynamic system involving multiple stakeholders such as ranchers, feedlots, processors, retailers, and consumers. The main stages of the U.S. beef supply chain are cow-calf operations, stocker/backgrounding operations, feedlots, meat packers and processors, and retailers. The production process begins with the cow-calf farms where cows and calves are raised. Following this stage, stocker/backgrounding operations take over, placing calves on grass or other types of roughage to promote growth. Feedlots are the final step of cattle production. At this step, cattle are given different rations of grain, silage, and/or protein supplements. Once finished, cattle are then sold to beef packers and processors where beef and beef by-products are produced and sold to retailers (USDA Economic Research Service, 2022). The economic size of beef-cattle industry including direct and indirect economic contributions during on-farm and post-farm activities was estimated as $167 billion in 2016 (English et al., 2020). Beef is also an indispensable nutrition in American dietary behaviors. It is a highly nutritious food as a source of protein, zinc, iron, and other minerals, B vitamins, and choline. Different types of beef have quality standards and are regulated and monitored by USDA. Beef quality grading is an important beef quality measure, and it is based on tenderness, juiciness, and flavor. There are eight quality grades in U.S. grading system: U.S. Prime, U.S. Choice, U.S. Select, U.S. Standard, U.S. Commercial, U.S. Utility, U.S. Cutter, and U.S. Canner. U.S. Prime has the highest and U.S. Canner has the lowest quality [1].

1.2 The Structure of the Dissertation

This dissertation includes three essays on the U.S. beef supply chain and aims to contribute to the current literature with its methods and results.

The first essay (Chapter 2) entitled "*Optimal Beef Cow Culling Strategies in the U.S.: A Dynamic Linear Programming Framework*" focuses on the economics of culling decisions in cow-calf operations in the U.S. Although culling decisions have been analyzed for decades, the literature has produced mixed results regarding the optimal culling age and the impact of culling decisions on herd profitability and productivity. This essay

[1] For further details for the facts and figures of the U.S. beef industry, see Erol (2022).

employs a novel application of a dynamic mathematical programming to construct a base model and conduct several experiments under various assumptions related to production and replacement heifer costs, cow fertility, calf weights, prices, and pregnancy check use. The objectives include assessing optimal culling decisions using the base model and experiments to determine under what conditions cows should be retained and estimating the impact of these decisions on herd structure and net returns. It also applies a pregnancy test experiment to analyze its impacts on the optimal culling decisions provided in the base model. Besides its methodological contributions in employing an infrequently used programming technique that is ideally suited to the problem considered, this study also has the potential to deliver practical guidance for beef cattle producers.

The second essay (Chapter 3) entitled "*Forecasting Beef-Cattle Prices in the Southern United States: A Model Selection Framework*" contributes to the U.S. beef cattle price forecasting literature with its model selection framework which compares traditional time series techniques and machine learning algorithms. The current U.S. livestock forecasting literature has utilized various techniques, including structural models, time series models, and machine learning approaches, to determine the best forecasting methods using both cash prices and basis. However, there are only a limited number of studies that have compared traditional time series models with machine learning models using cash prices. This essay applies ARIMA-type models with and without seasonal components and machine learning techniques in a rolling origin cross validation scheme to select the best technique to provide one-week-ahead price forecasts for steer, heifer, and cull cow cash prices in the southern U.S. It provides a detailed discussion of selected machine learning models (ridge regression, LASSO regression, random forests, and gradient boosted machine) and cross validation strategy. It also compares the prediction performance of the methods in small and large samples. The study performs these techniques using weekly Kentucky cattle auction prices with lagged variables and dummy variables for weekly seasonal structure. The approaches and outcomes of the essay offer valuable tools for livestock price forecasting, which can be utilized by extension specialists and producers to enhance their decision-making processes.

The third essay (Chapter 4) entitled "*The COVID-19 Shock and Dynamics of Price Adjustment in the U.S. Beef Sector*" examines the price dynamics along the U.S. beef supply chain and analyzes the impact of COVID-19 on the dynamics of vertical price transmission within the U.S. beef industry using monthly farm, wholesale, and retail prices. Although numerous studies have examined the price transmission dynamics in the vertical chain of U.S. beef markets, only a limited number of studies have measured the impact of external shocks on these dynamics. In this essay, a vertical error correction model is utilized with historical decomposition graphs to estimate the impact of the pandemic on price adjustment. The methodology used in this essay accounts for endogenous structural breaks in the long-run cointegration relations of price series for the period from January 1970 to April 2021 and, to the best of my knowledge, it is only study that estimates the impact of COVID-19 on all stages of the beef supply chain, including the recent periods. The empirical results of the essay provide several implications to contribute to the current literature and construct a base for policy reactions for similar crises.

The final chapter (Chapter 5) concludes the dissertation with the empirical results and discussions of each chapter.

Tables and Figures for Introduction

Table 1.1 U.S. Beef Production

Year	Beef Production (Million Pounds)
2020	27,243
2019	27,224
2018	26,939
2017	26,251
2016	25,288
2015	23,760
2014	24,315
2013	25,790
2012	25,989
2011	26,270

Data Source: USDA, National Agricultural Statistics Service.

Table 1.2 U.S. Beef-Cattle Inventory

Year	Number of Head (Thousands)
2020	31,339
2019	31,691
2018	31,466
2017	31,171
2016	30,164
2015	29,332
2014	28,956
2013	29,631
2012	30,282
2011	30,913

Data Source: USDA, National Agricultural Statistics Service.

Table 1.3 U.S. Beef-Cattle Inventory by State-2020

Year	Number of Head (Thousands)
TEXAS	4,570
OKLAHOMA	2,109
MISSOURI	2,083
NEBRASKA	1,922
SOUTH DAKOTA	1,783
KANSAS	1,443
MONTANA	1,428
KENTUCKY	1,021
NORTH DAKOTA	995
ARKANSAS	915
U.S. Total	31,339

Data Source: USDA, National Agricultural Statistics Service.

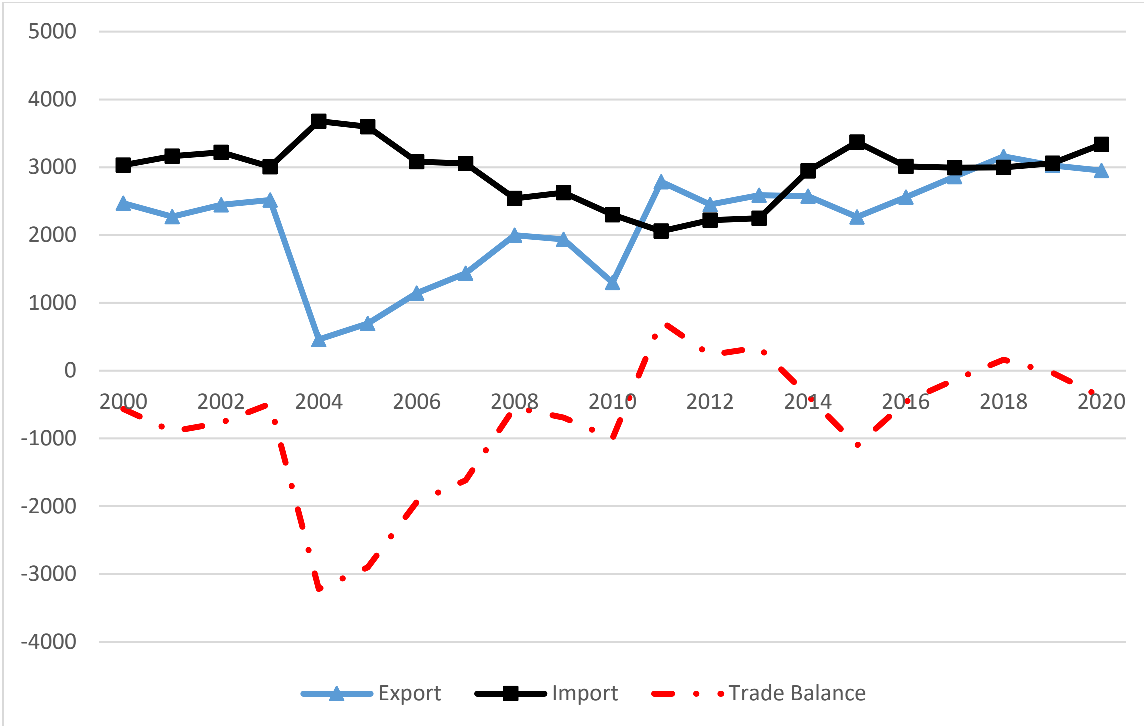

Figure 1.1 U.S. Beef and Veal Trade (Million Pounds)
Data Source: USDA, Economic Research Service, Livestock and Meat Trade Data.

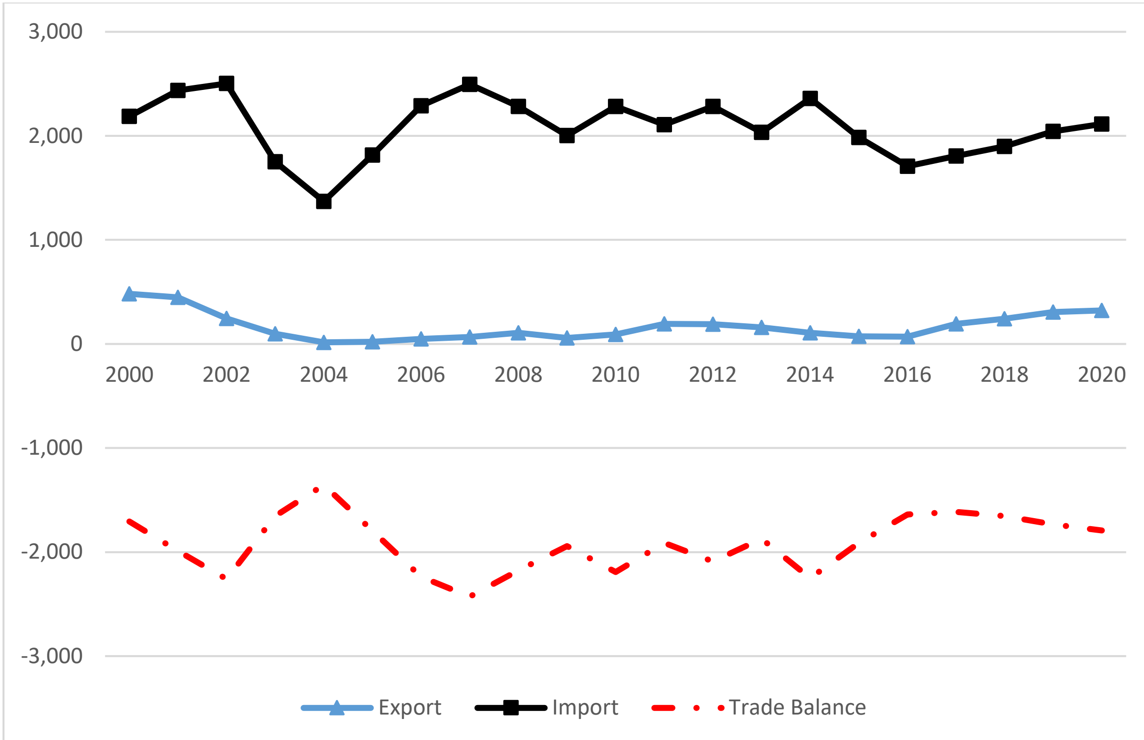

Figure 1.2 U.S. Cattle Trade (Thousand Head)
Data Source: USDA, Economic Research Service, Livestock and Meat Trade Data.

CHAPTER 2. Optimal Beef Cow Culling Strategies in the U.S.: A Dynamic Linear Programming Framework

2.1 Introduction

Beef cow culling is defined as removing selected beef cows from the herd permanently. Culling decisions, i.e., the selection and timing of cows to cull, are an essential part of farm management to sustain profitability and productivity of the herd in the short run and long run. According to USDA data, the culling rate (the percentage of cows removed permanently from herd each year) was 12.9% in all operations and 18% in small operations where herd size is 1-49 cows in 2017 (USDA National Animal Health Monitoring System, 2020). Studies estimate the revenue generated from cull cow sales to be about 15-30% of a yearly revenue (Amadou et al., 2014; Blevins, 2009; National Cattlemen's Beef Association, 2016). The reasons for cow culling are related to both biology and economics. Reproductive efficiency including pregnancy status and other fertility problems, age, disposition, cow's health and physical soundness, concerns of producing inferior calves, and a desire for genetic improvement from replacement breeding stock are primary biological reasons to cull a cow (Arnold et al., 2021; Hersom et al., 2018). In addition to the cow health and herd structure, cow-calf prices and seasonal trends in the markets, production costs including maintenance and replacement heifer costs, expected future earnings, cash flow, and risk management are major economic factors affecting culling decisions (Hersom et al., 2018; Peel & Doye, 2017; Ward & Powell, 2017). Therefore, ranchers and farm managers can improve both profitability and productivity of the herd by utilizing a data-driven and informed culling strategy.

There is an extensive literature analyzing the reasons and motivation behind beef cow culling and evaluating optimal culling strategies under a variety of assumptions

(Azzam & Azzam, 1991; Bentley et al., 1976; Boyer et al., 2020; Frasier & Pfeiffer, 1994; Ibendahl et al., 2004; Mackay et al., 2004; Melton, 1980; Tronstad & Gum, 1994). As parallel to some of these study's conclusions, culling open cows immediately after pregnancy check is advised in practice where pregnancy check is not common (USDA National Animal Health Monitoring System, 2020). However, the literature has mixed conclusions and underlines the impact of fertility rates, age of open cows, prices, and costs on the optimal culling strategies that may improve herd productivity and provide a flexible strategy to cope with price cycles and costs.

This study develops a model framework using dynamic linear programming for cow-calf operations in the U.S. to determine optimal beef cow culling age and measuring the sensitivity of the optimal strategy with respect to other factors such as cow's fertility, cow-calf prices, variable costs, and replacement prices. The model is solved with the data obtained for a spring calving herd that sold weaned calves in the fall in Kentucky. In addition to the methodological contributions, this research aims to provide practical guidance for beef cattle producers.

2.2 Literature Review

The economics of culling decisions on cow-calf operations in the U.S. has been analyzed considering various aspects in the literature for decades. The majority of studies that specifically focus on the optimal culling age employed a net present value framework by either comparing the opportunity cost between retaining a cull cow and replacing with a bred heifer, or evaluating the contribution of cow to herd future revenue streams throughout the productive years of the cows (Bentley et al., 1976; Boyer et al., 2020; Ibendahl et al., 2004; Mackay et al., 2004; Melton, 1980; Trapp, 1986). These studies based their models on variations of asset replacement methodology discussed in Burt (1965) and Perrin (1972). Azzam and Azzam (1991) and Frasier and Pfeiffer (1994) worked with Markovian multi-stage decision analysis with transition probabilities of different states including cow's age, productivity, calving season, body condition scores, and calving dates. Dynamic programming has also been applied to find the optimal decision rule to cull a cow (Tronstad & Gum, 1994).

The literature on culling versus keeping open cows has been mixed. Several studies have suggested that all open cows should be culled. In their Markovian decision analysis, Azzam and Azzam (1991) used cow's productive status (open, pregnant, and unsound) in combination with cow's age (2.5 to 10.5) and two calving seasons (spring and fall) for Nebraska as transition states and two decisions (keep and replace). They recommended replacing all open cows with spring-born heifers and retaining any pregnant cows of any age. Using a similar methodology augmented with the impacts of herd management practices, Frasier and Pfeiffer (1994) also suggested culling all open and late lactating cows. Their sensitivity analysis underlined the impact of different cow and calf prices and replacement heifer cost on the optimal culling strategies. The most recent study on the optimal culling decisions, Boyer, Griffith, and DeLong (2020) worked with Tennessee herd level data for both spring and fall calving seasons and analyzed the impact of productivity failure (missing a calf during the production life) on the operation's long-term profitability. They concluded that selling an open cow after missing one calf would be a better option to increase the profitability instead of keeping and rebreeding the cow.

The literature that has left the door open to retention has focused on the calving system, varying fertility of cow across ages, genetic improvement, the importance of prices, and replacement costs in this determination. Bentley, Waters, and Shumway (1976) found the optimal replacement policy to be replacing a cow after her seventh calf is sold since the expected present value of the cow is maximized at age 8. In their sensitivity analysis, they conclude that changing cattle prices and feed costs does not impact the optimal policy. They demonstrated that culling decisions are sensitive to lower calving rates and changing cull cow prices with carcass quality (i.e., lower prices for older cows), which lead to an earlier optimal culling. Melton (1980) investigated the impact of an endogenous genetic progress in the herd on the optimal culling age with experimental herd data in Florida. Results suggested a culling age of 8 under genetic improvement and 11 without genetic progress. Tronstad and Gum (1994) performed a stochastic dynamic programming model with biannual calving, cow fertility estimates, and stochastic prices under a multi-period horizon with the objective of maximizing expected wealth. Their model results emphasized that a dual calving system would be more profitable for ranchers. They also utilized a Classification and Regression Trees (CART) technique to provide more practical and interpretive culling advice based on dynamic programming model results. Their decision trees included "keep" and "replace" decisions and were based on 1,100 combinations of state variables (cow age, pregnancy status, calf price, replacement price, and cull cow prices) used in the model. Their tree-based analysis identified age for open cows and calf prices for pregnant cows as the most important variable in culling. CART results also rejected the culling strategy which suggests culling open cows all the time regardless of their productivity and stated that when spring and fall calving are possible that would make rebreeding possible, open cows should be kept 26% of time. This rate was found as 95% for pregnant cows in the study.

One of the early studies evaluating the impact of the prices on the culling decisions and herd decomposition, Trapp (1986) assumed a varying herd size and nonconstant prices in a simulated herd model and pointed out a flexible culling strategy with culling ages varying from 5 to 12 to manage cyclical prices. Mackay et al. (2004) discussed the impact of prices on marketing strategies based on analysis using data from a Nebraska cow-calf operation. They pointed to the changing herd decomposition under different prices; an older herd was better when prices were lower, and a younger herd was better when prices were higher. Ibendahl, Anderson, and Anderson (2004) studied the optimal culling policy within a net present value framework with fertility data obtained from Tronstad et al. (1993). They contradicted the conclusion of the studies which has been to cull open cows by focusing on the age of open cows and suggested a flexible culling strategy to deal with production costs. They emphasized the impact of replacement and production costs on the replacement decisions and recommended retaining younger open cows when calf crop value and production costs are low and the difference between replacement heifer cost and cull value is high.

The culling rate is about 30% in dairy operations (USDA National Animal Health Monitoring System, 2020) and culling strategies are examined with a focus on the cow performance and milk production in the dairy literature (Cabrera, 2010; Lehenbauer & Oltjen, 1998; van Arendok & Dijkhuizen, 1985; van Arendonk, 1986). Van Arendok and Dijkhuizen (1985) studied optimal policies for open cows with a variation in the time of conception and three alternatives: inseminating, leaving her open, and cull her immediately. They underlined the impact of replacement costs on culling open cows. Cabrera (2010) applied a Markovian linear programming model with a net revenue maximization objective and solved the model with "keep" and "replace" decisions under different dietary treatments and state variables which are defined by parity, month in lactation, and pregnancy status. The model suggests that pregnant cows should be kept regardless of their production performance and higher culling rates should be allowed when milk prices and replacement costs are low and corn prices are high. Open cows are culled earlier depending on market conditions.

This study aims to build upon this literature by constructing a model framework for cow-calf operations in the U.S. to evaluate culling strategies. The objectives of the study are to evaluate optimal culling decisions with a base model and several experiments performed with a variety of assumptions related to cost, fertility, weights, and prices, determine under what conditions open cows should be retained, and estimate the impact of pregnancy checking on the herd profitability and culling decisions. In addition to its methodological contributions in applying an infrequently used programming technique ideally suited to the problem considered, it also has potential to provide practical guidance for beef cattle producers.

2.3 Methodology and Data

The economic decision-making framework of a commercial size beef cow-calf producer is formulated as a single year dynamic linear programming model since cow culling decisions depend on the dynamic nature of changes in productivity by age of brood cow and stochastic factors (probabilities of pregnancy, cow death, calf loss, etc.). Optimal decision modeling of problems with such dynamic elements can be formulated with both multi-period and single period dynamic linear programming models. In multi-period models, the number of periods can be assumed to be known or unknown depending on the problem's nature and research considerations and optimal solutions are generated with the assumption of disequilibrium where decisions vary over a number of time periods. Single period dynamic linear programming models which are also called equilibrium models, assume that optimal decisions are repeatedly made in all time periods and defined as long run steady state solutions (McCarl & Spreen, 1997). The model used in the study is a single year steady-state equilibrium of unknown asset life with prices, costs, and probability values that is solved with results then being interpreted. The operation's size is assumed constant by imposing a herd size constraint in the model formulation.

The dynamic linear programming model used in the study is specified as:

Max: *Net Return above Selected Costs*

$$\sum_{Y\$(ORD(Y)\leq 12)}\Big(\sum_{P} Prices(P,Y) * Sale(P,Y) - \sum_{C} Production\ Cost(C,Y) * Production(C,Y)\Big) \tag{2.1}$$

Subject to:

Herd Size:

$$\sum_{(C,Y)\$(ORD(Y)\leq 12)} Production(C,Y) \leq 100 \tag{2.2}$$

Market Balance

$$Sale(P,Y) - \sum_{C} Fertility\ Probability(C,P,Y) * Production(C,Y) \leq 0 \quad \forall\ P,Y \tag{2.3}$$

Linkage between ages:

Pregnant:

$$\sum_{(C,Y)\$(ORD(Y)<12)}\big(Retention\ Probability(C,'Pregnant',Y) * Production(C,Y)\big) - Production('Dam',Y+1) \leq 0 \quad \forall\ Y < 12 \tag{2.4}$$

Open:

$$\sum_{(C,Y)\$(ORD(Y)<12)}\big(Retention\ Probability(C,'Open',Y) * Production(C,Y)\big) - Production('Open',Y+1) \leq 0 \quad \forall\ Y < 12 \tag{2.5}$$

Cull:

$$\sum_{(C,Y)\$(ORD(Y)<12)} \left(Fertility\ Probability(C,'\ Cull', Y) * Production(C,Y)\right) - \sum_C Production(C, Y+1) - Sale('Cull', Y) \leq 0 \quad \forall\, Y < 12 \quad (2.6)$$

Where: C is the index of 3 different types of cows depending on their pregnancy status: cow with calf at side from previous year (dam), cows with no calf from previous year (open), and first-time heifer (two-year-old bred heifers). Two-year-old bred heifers are purchased to replace cull cows. P denotes the 3 different products which are being sold at the market: steer, heifer, and cull cows. Y is the age (2 to 13 years old herein) of each cow. $Production(C,Y)$ is the number of cows that are raised in the operation at the equilibrium solution (repeated optimal decisions as discussed above). $Sale(P,Y)$ is the number of calves and cull cows that are sold under the equilibrium solution. $Production\ Cost(C,Y)$ includes annual variable cost per cow and ownership costs of a bred replacement heifer. There are two probability series in the model: fertility and retention.

Fertility Probabilities(C,P,Y) are the chance of products P (having a steer, heifer, and live cow at year end) based on age Y and prior year status C (cow with calf at side or open last year). Retention Probabilities(C,K,Y) reflect the percentage of cows that have a calf at side or are open this year (after including the chance of her death or being unsound) and are available for retention to next year. Equation 2.1, the objective function is to maximize the expected herd level net return above selected costs while equations 2.2–2.6 impose various constraints. A commercial operation herd size is assumed with a maximum of 100 cows allowed in equation 2.2. The herd size of 100 is selected to formulate a medium scale cow-calf operation (50 to 199 cows) and make results more interpretable in terms of percentages. Equation 2.3 assures market balance by limiting sales to the amounts produced for every cow age. Linkage equations (2.4-2.6) are age sequencing constraints which ensure that the number of cows C of age Y+1 must be less than or equal to number of that cow type that were kept until age Y in a standard dynamic linear program fashion.

Fertility and Retention probabilities given in Table 2.1 are calculated based on calving rates and fertility estimates obtained from Tronstad et al. (1993). The probability of calving, survival, and retention vary depending on cow's age and her productivity status in the previous year. For example, if a cow is 5 years old and had a calf at her side last year, she has a 76.5% chance of having a calf this year and a 96.6% chance of surviving after calving. The probability of this live cow calving again is 71.9% and the probability of her failing to calve is 22.0% (the first panel of Table 2.1). The fertility rates are lower for cows without calves from previous year (the second panel of Table 2.1). These cows are assumed either have lost their calf or failed to calve in the previous season. On the other hand, if a cow of the same age did not have a calf at side last year, her chance of calving this year is 64.7% and her chance of surviving increases to 98.3%. She also has a lower probability (58.3%) to calve this year and higher probability (31.8%) to fail to calve again. The calving rate is assumed to be 97.8% for replacement heifers. Calf survival rate from birth to weaning is assumed to be 95.5% for all cows with 4.5% calf death loss (Strohbehn, 1994).

The price data is obtained from the USDA Agricultural Marketing Service (AMS) and is given in Table 2.2. The 10-year (2012-2021) arithmetic average of monthly October and November feeder steer, heifer, and cull cow prices were used in the study to account for fall sales for spring calving herds in Kentucky. Steer and heifer prices for medium and large frame size and #1-2 muscling calves are utilized. Prices per pound are also adjusted downward to account for the heavier calves (commonly referred as price slide). Steers and heifers from 5- to 10-year- old dams are assumed to be weaned and sold at an average weight of 600 and 550 lb respectively. The weights for other calves are adjusted based on dam's age with Beef Improvement Federation (2018) data. Cull cow prices are estimated using historical USDA-AMS price data from the 80-85% boning cow category. Cull cow price per lb is assumed to be \$0.62 per lb with an average cull cow weight of 1,200 lb. To adjust cull cow prices with respect to carcass quality across ages, breaking grade cull cow prices (for age 2) and lean grade cull cow prices (for age 13) are used as minimum and maximum prices and adjusted price data is computed by decreasing prices linearly from age 2 to age 13.

Annual cost per cow by age is computed for a spring calving herd and it covers annual variable costs, cow depreciation, interest costs, and bred replacement heifer prices. Annual variable costs are obtained from Halich, Burdine, and Shepherd (2022) and are shown in Table 2.3. These estimates are made for a spring calving cow-calf operation in 2021 and include only cash costs for the operation. The pasture stocking rate is assumed to be 2.0 acres per cow and hay consumption is assumed to be 2.5 tons per cow. The operation has its own pastureland and produces its own hay. Since operation costs presented in Table 2.3 may vary by herd size and management, an annual variable cost experiment is implemented to account for different annual cost per cow values and their impact on optimal culling strategies. Breeding costs are excluded from replacement heifers' annual variable cost as they are purchased already bred. Annual depreciation and interest are computed with 3% interest rate for a \$1,500 bred heifer which has 11 productive years and \$700 cull cow salvage value.

The model timeline constructed for the spring calving herd based on common practices in the South is presented in Figure 2.1. The model starts in October at production year t after all calves were weaned and sold in weaning season of production year t-1 and ends in September. Therefore, there are no calves from the previous calving season during the model period and calf crops born are sold in the following weaning season.

The base model is a single year steady-state equilibrium of unknown asset life with base prices, costs, and probability values that is solved with results then being interpreted. Various experiments are conducted to analyze the study objectives by resolving the model after changing relevant coefficient values. This allows for the evaluation of the robustness of the model and performance of a sensitivity analysis by comparing experiment results to those of the base model.

2.4 Results and Discussion

The base model results can be seen in tables 2.4 and 2.5 and suggest that producers should cull all cows that are older than 10 based on their productivity, production costs, and price assumptions made. This result reflects the reproductive performance of the cows that is maximized between age 4 to 9 and starts to decline after age 10 (Arnold et al., 2021; Ward & Powell, 2017). The operation makes a modest net return above specified costs of \$12,347 and purchases 15.6 bred replacement heifers annually. The 100-cow operation consists of 70.7 cows with calves at their side, 15.6 first-time heifers, and 13.8 open cows. The model suggests culling open cows earlier (at age 7) than cows with calf at side (at age 10) given their productivity status and probabilities. Each year, the operation sells 36.8 steers, 36.8 heifers, and 12.6 cull cows in the base model. The average cow age in the herd is 5 years and the average age of cows culled is 7.7 years with a 12.6% culling rate in the base model. These results are consistent with the literature which suggests the optimal culling age between 7 and 11 (Bentley et al., 1976; Mackay et al., 2004; Melton, 1980; Trapp, 1986; Tronstad & Gum, 1994).

2.4.1 Sensitivity Analysis

To measure the sensitivity of the optimal results, several experiments are run with different prices, costs, fertility values, and genetic improvement resulting in weight gain. A pregnancy check experiment is also run to estimate the impact of pregnancy checking on the returns. The experiments' outcomes underline the sensitivity of the optimal strategies to market conditions, cost structure, cow fertility, calf weights, and pregnancy check use (Table 2.4 – 2.6). That being said, optimal culling age and herd structure do not change under the experiments with adjusted cull cow prices by age (to account for decreasing carcass quality as cows get older) and simultaneous price changes in calf and replacement heifer prices (to reflect the price transmission between calf and replacement heifer markets). Changing the annual variable cost assumption also does not impact the optimal culling age, production level, or sale amount. This was an expected outcome since annual variable costs are assumed to be same across cow ages. These three experiments cause changes only in net return above selected costs comparing to base model results.

2.4.2 Prices

To evaluate the impact of different price assumptions on the optimal culling strategies, all output prices were changed by 10%. As expected, a simultaneous 10% increase in cow and calf prices enhances steer and heifer calf values, leading to a suggestion that producers should cull open cows older than 6 (one year sooner than the base model) and cull all cows older than 9. As shown in Table 2.5, calf sales rise to 74.8, the number of cows with calf at side increases to 71.6, and the number of open cows decreases to 11.3. On the other hand, a 10% decrease in prices favors a strategy of culling later at age 8 for open cows. The optimal culling age for all cows is the same as the base model but the production amount is higher compared to the base model. This is a rational behavior since producers can take advantage of good market conditions and sell their calves and less productive cows, specifically by culling earlier when prices are high and keeping their cows and calves when prices are low. These results are consistent with Mackay et al. (2004) and Trapp (1986) who pointed to the varying herd decomposition with prices.

To examine the impact of cull cow prices on decision variables, the model is run with a 25% increase only in cull cow prices. The results show a 20.8% increase in expected net return and the optimal strategy becomes culling cows with calves at their side at age 8 and open cows at age 6. As expected, a stronger cull cow market encourages an earlier culling strategy. The experiment with the same amount of decrease in cull cow prices suggests a later culling strategy in which the optimal culling age is 8 for open cows and 11 for all cows. These results underline the correlation between cull cow prices and culling age.

2.4.3 Costs

The model includes three cost components: variable production costs, replacement heifer purchase price, and interest costs. Cost experiments for each component are performed separately to assess their individual impacts on culling decisions. Although the experiment with annual variable costs does not alter optimal culling decisions, the optimal culling strategy in the base model is highly sensitive to ownership costs of bred replacement heifer which includes both heifer value and interest rate. Replacement heifer costs are a major part of operation costs for those producers who prefer to purchase a replacement heifer instead of raising it in the operation (Halich et al., 2022). When replacement is costly, producers tend to keep cows longer despite their declining productivity.

A 20 % increase in replacement heifer value leads to a culling age of 11 for cows with calf at side and 9 for open cows. The same percentage decrease in replacement heifer values suggests an optimal culling of cows with calves at their side at age 8 and open cows at age 5. These are the largest changes in optimal age among all experiments performed. The average cow age in the herd is 4.3 years and the average age of cows culled is 6.4 years with a 16.6% culling rate when replacement heifer prices decrease by 20%. Producers buy 13.9 bred heifers when replacement heifer value increases by 20% and 19.1 replacement heifers when replacement heifer value is 20% cheaper compared to the value in base model.

Interest rate changes also affect results by changing optimal culling age, net return, and herd structure. A 12% interest rate which makes replacement heifer purchases more costly for the producer, results in a later culling age of 8 for open cows and age 11 for all cows. The herd becomes older with an average cow age of 5.3 years.

2.4.4 Fertility

Cow's genetics, body condition, and age are major determinants of reproductive efficiency and can be improved by appropriate management practices (Corah & Lusby, 2000; Tronstad et al., 1993).

The selection of better replacement heifers, improvement of health and nutrition programs, and adoption of advanced technologies in fertility assessment and improvement can lead to more productive beef cows in the herd (Moorey and Biase 2020). The model is also solved under different productivity values to see the impacts of enhanced management on optimal culling decisions. To this end, the probability of an open cow to be pregnant in the next year is changed by 1% across all ages, and calving rate is changed by 1% for age 2-12 and 0.4% for age 13. An increase in calving rate and fertility rates generate an additional 11.3% net return above specified costs and a similar optimal culling strategy of production with same ages but higher sale amounts and culling rates compared to the base model's outcomes.

2.4.5 Weight Gain

An experiment is also run to measure the impact of calf weights on the optimal culling strategy. Steer and heifer weights are increased by 20% over the base model. Prices per pound are also adjusted downward to account for the heavier calves. A 20% increase in calf weights results in a considerable change in net return with a younger herd and an early culling strategy. The optimal culling age moves to 6 for open cows and 9 for dams. The operation sells 37.4 steers, 37.4 heifers, and 14.3 cull cows.

2.4.6 Pregnancy Check

Although pregnancy checking is generally advised, it is not common among U.S. beef operations. According to USDA National Animal Health Monitoring System (2020) data, the percentage of operations that regularly pregnancy checking their cows (palpation, blood test, and ultrasound) was 31.6 for all operations in 2017. The reasons are labor and time costs, test costs, and producers' beliefs and habits. To estimate the impact of pregnancy checking on the optimal culling decisions, an experiment with pregnancy checking is run. In the experiment, the marginal cost of pregnancy check is assumed to be $10 per cow with $150 trip charge and annual variable costs in the base model are adjusted accordingly. Fertility data is recalculated for the experiment.

The base model, as defined by equations 2.1-2.6 is modified to run the pregnancy checking use experiment. The model objective function, herd size, market balance, and cull cows' linkage constraints are the same as in the base model. Pregnant and open cows' linkage constraints are adjusted, and two additional constraints are added to the model:

Linkage between ages:

Pregnant:

$$\sum_{(C,Y)\$(ORD(Y)<12)} \big(Retention\ Probability(C,'Pregnant',Y) * Production(C,Y)\big) - Production('DamTestPregnant',Y+1) - Production('DamTestNonpregnant',Y+1) \leq 0 \qquad \forall\, Y < 12 \tag{2.7}$$

Open:

$$\sum_{(C,Y)\$(ORD(Y)<12)} \big(Retention\ Probability(C,'Open',Y) * Production(C,Y)\big) - Production('OpenTestPregnant',Y+1) - Production('OpenTestNonpregnant',Y+1) \leq 0 \qquad \forall\, Y < 12 \tag{2.8}$$

Pregnancy Test Constraints:

Dam:

$$Production('DamTestPregnant',Y) * PregConstProbabilities('DamTestNonpregnant','DamPregRateConst',Y) - Production('DamTestNonpregnant',Y) * PregConstProbabilities('DamTestPregnant','DamPregRateConst',Y) - Sale('Cull',Y) * PregConstProbabilities('DamTestPregnant','DamPregRateConst',Y) \leq 0 \tag{2.9}$$

Open:

$$Production('OpenTestPregnant', Y) * PregConstProbabilities('OpenTestNonpregnant', 'OpenPregRateConst', Y) - Production('OpenTestNonpregnant', Y) * PregConstProbabilities('OpenTestPregnant', 'OpenPregRateConst', Y) - Sale('Cull', Y) * PregConstProbabilities('OpenTestPregnant', 'OpenPregRateConst', Y) \leq 0 \quad (2.10)$$

There are 5 different types of cows (index C) in the pregnancy check use model: dam tested as pregnant, dam tested as non-pregnant, open tested as pregnant, open tested as non-pregnant, and first-time heifer. Two-year-old bred heifers are not tested since they are purchased as bred heifers. The linkage equations (2.7-2.8) are modified versions of equation 2.4 and 2.5 and ensure that the number of cows C of age Y+1 must be less than or equal to number of that cow type that was kept until age Y. Equations 2.9 and 2.10 are the pregnancy test constraints and balance the production and cull amounts after testing dam and open cows for every cow age to ensure proper ratios of pregnant cows. The cow is tested as pregnant or non-pregnant and culled after the pregnancy test.

The results are presented in Table 2.6. The model results suggest that producers should keep dam cows that are tested as pregnant until age 13 and cull all open cows that are tested as non-pregnant. This indicates that a producer is better off starting with a cow that is already confirmed to be pregnant each year, in this case a bred heifer. Based on the herd decomposition following pregnancy checking, the model only keeps open cows that are 4-year-old and younger and tested as pregnant. These results suggest that the producers should consider keeping younger open cows and contradict the common practice in the field and the studies that suggest selling open cows under any circumstance (Azzam & Azzam, 1991; Boyer et al., 2020; Frasier & Pfeiffer, 1994). The 100-cow operation consists of 73.1 cows with calves at their side tested as pregnant, 2.9 cows with calves at their side tested as non-pregnant, 3.3 open cows tested as pregnant, and 20.7 first time heifers. The average cow age in the herd is 4.24 years and the average age of cows culled is 6 years. The net return above selected costs increases by 68.2% compared to the base model because of the culling management changes enabled by pregnancy check information and increased calf sales. Each year, the operation sells 44.9 steers (about 8 more than the base model), 44.9 heifers, and 17.3 cull cows (4.7 more than the base model) in the model.

2.5 Conclusion

Developing a culling strategy has great influence on financial and structural soundness of cow-calf operations and draws substantial attention from academics and extension field specialists. This study contributes to the current discussions with its methodology and results.

A single year dynamic linear programming model is formulated and run with cow fertility estimates as well as price and cost data obtained for a spring calving herd in Kentucky to provide a set of culling strategies with optimal cow culling age for beef cattle producers. The results of the base model suggest that producers should cull all cows older than age 10 and all cows that fail to calf once they reach the age of 7. Given the base cost and price values, the 100-cow operation generates a net return above selected costs of $12,347 and produces 70.7 cows with calves at their side, 15.5 bred replacement heifers and 13.8 open cows. Each year, the operation sells 73.6 calves and 12.6 cull cows. The average cow age in the herd is 5 years and the average age of cows culled is 7.7 years with a 12.6% culling rate in the base model.

The impact of each variable on the optimal decisions is also measured with several experiments to evaluate the sensitivity of the optimal strategy to changes in the markets, farm production costs, and management practices. The outcome of these experiments underlines the sensitivity of the optimal strategies to market conditions, particularly calf/cull cow prices, ownership costs of bred replacement heifer, and herd management skills to increase cow's productivity and calf weights as essential parameters in optimal culling. While cow-calf price changes impact net return values to a considerable extent among experiments conducted, the cost sensitivity analysis with changing bred heifer replacement value alter both net return and herd age decomposition most substantially. Culling age in the base model (age 7 for open cows and 10 for cows with calf at their side) decreases to 5 for open cows and 8 for cows with calf at their side when replacement prices decrease by 20% and increase to 9 for open cows and 11 for cows with calf at their side when replacement prices increase by 20%.

When pregnancy checking is incorporated into the model, net return above selected costs increase by 68.2%. Producers only keep open cows that are 4-year-old and younger. Previous literature has been mixed on the retain/replace decision for open cows and this work suggests that younger open cows should be kept. It is also worthwhile to note that the assumed cost of pregnancy checking is relatively low at $10 per cow, which largely assumes that cattle were already being worked.

One of the primary implications of the sensitivity analysis is that producers should pay a considerable amount of attention to management practices to monitor and improve cow's productivity and calf weights since better management creates a potential to have lower number of open cows in the operation and earn higher net returns. The impact of higher weaning weights and increased probability of weaning calves results in substantive improvement in returns above the base model suggestions.

Tables and Figures for Chapter 2

Table 2.1 Fertility and Retention Probabilities

	Cows with Calf from Previous Year											
	Cow Age (Year)	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
Fertility	% Calving	79.7	78.2	76.5	74.7	72.6	70.4	67.9	65.2	62.2	59.0	55.5
	% Cow survival	97.9	97.3	96.6	95.9	95.2	94.5	93.8	93.1	92.4	91.8	91.1
Retent.	% Survival, sound, and back-to-back calves	76.9	74.6	71.9	68.8	65.3	61.6	57.5	53.2	48.7	43.9	39.0
	% Survival, sound, but no calf this year	19.6	20.8	22.0	23.3	24.6	25.9	27.2	28.4	29.5	30.5	31.3
	Cows without Calf from Previous Year											
	Cow Age (Year)	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
Fertility	% Calving	69.0	67.0	64.7	62.0	59.3	56.4	53.7	51.0	48.7	46.7	45.2
	% Cow survival	99.4	98.9	98.3	97.6	96.7	95.7	94.6	93.3	91.9	90.3	88.7
Retent.	% Survival, sound, and calf this year	66.3	62.6	58.3	53.7	49.0	44.3	40.0	36.0	32.5	29.6	27.4
	% Survival, sound, but no calf again	29.8	30.8	31.8	32.8	33.7	34.2	34.5	34.6	34.3	33.8	33.2

Notes: Data in the table are calculated based on calving rates and fertility estimates obtained from Tronstad et al. (1993).

Table 2.2 Calf Weights and Prices

Cow Age	Steer		Heifer	
	Weight(lb)	Price($/lb)	Weight(lb)	Price($/lb)
2	540	1.60	496	1.48
3	560	1.59	514	1.47
4	580	1.57	532	1.46
5	600	1.55	550	1.44
6	600	1.55	550	1.44
7	600	1.55	550	1.44
8	600	1.55	550	1.44
9	600	1.55	550	1.44
10	600	1.55	550	1.44
11	580	1.57	532	1.46
12	580	1.57	532	1.46
13	580	1.57	532	1.46

Notes: Author's calculations based on data obtained from USDA Agricultural Marketing Service.

Table 2.3 Annual Variable Costs

Item	Annual Cost ($)
Pasture Maintenance (Cash costs:2.0 acres and $20 per acre)	40
Hay (Cash costs: 2.5 tons and $35 per ton)	88
Mineral	35
Vet	25
Breeding	40
Marketing	25
Winter Feeding and Other Machinery (cash costs)	15
Trucking (calves, supplies, etc.)	15
Others (insurance, property taxes, water, etc.)	40
Total	**323**

Source: Halich, Burdine, and Shepherd (2022)

Table 2.4 Optimal Culling Strategies: Net Return above Selected Costs and Culling Age

					Herd Net Return (100 cows)		Cull Age		Weighted Average	
	Price	Costs	Interest Rate	Fertility	$	% Change[c]	Open	All	Cow Age	Cull Age
Base Case	Base[a]	Base	Base[b]	Base	12,347	-	7	10	4.97	7.7
				Experiments						
Price	+10% All	Base	Base	Base	19,567	58.5%	6	9	4.65	7.07
	-10% All	Base	Base	Base	5,272	-57.3%	8	10	5.16	8.14
	+25% Cull	Base	Base	Base	14,909	20.8%	6	8	4.48	6.84
	-25% Cull	Base	Base	Base	10,096	-18.2%	8	11	5.26	8.25
Cost	Base	+20% Bred Heifer	Base	Base	7,236	-41.4%	9	11	5.44	8.63
	Base	-20% Bred Heifer	Base	Base	18,208	47.5%	5	8	4.31	6.36
	Base	Base	12%	Base	848	-93.1%	8	11	5.26	8.25
	Base	Base	0.01%	Base	16,187	31.1%	7	9	4.84	7.54
Productivity	Base	Base	Base	+1%[d]	13,747	11.3%	7	10	4.97	7.69
	Base	Base	Base	-1%[d]	10,960	-11.2%	7	10	4.97	7.73
20% Weight Gain	Adjusted[e]	Base	Base	Base	20,162	63.3%	6	9	4.65	7.07

Notes: [a]Steer and heifer prices are age adjusted and cull prices are same for all ages, [b]Interest rate is 3% in the base model, [c]Change from net return in the base case, [d]Pregnant, live calf: 1% for age 2-12 and 0.4% for age 13 & open to pregnant: 1%, [e]Steer and heifer prices are adjusted.

Table 2.5 Optimal Culling Strategies: Production and Sales

					Head Produced		Replacements Bought	Head Sold	
	Price	Costs	Interest Rate	Fertility	Dam[c]	Open	Bred Heifers	Steers/Heifers	Cull Cows
Base Case	Base[a]	Base	Base[b]	Base	70.69	13.76	15.55	36.8	12.58
					Experiments				
Price	+10% All	Base	Base	Base	71.65	11.32	17.03	37.4	14.27
	-10% All	Base	Base	Base	69.14	16.13	14.73	36.4	11.64
	+25% Cull	Base	Base	Base	70.53	11.76	17.70	37.6	15.07
	-25% Cull	Base	Base	Base	69.66	15.86	14.48	36.3	11.32
Cost	Base	+20% Bred Heifer	Base	Base	68.29	17.85	13.86	35.9	10.56
	Base	-20% Bred Heifer	Base	Base	72.43	8.43	19.13	38.0	16.58
	Base	Base	12%	Base	69.66	15.86	14.48	36.3	11.32
	Base	Base	0.01%	Base	69.92	14.12	15.96	36.9	13.09
Productivity	Base	Base	Base	+1%[d]	70.69	13.76	15.55	37.6	12.70
	Base	Base	Base	-1%[d]	70.69	13.76	15.55	36.0	12.45
20% Weight Gain	Adjusted[e]	Base	Base	Base	71.65	11.32	17.03	37.4	14.27

Notes: [a]Steer and heifer prices are age adjusted and cull prices are same for all ages, [b]Interest rate is 3% in the base model, [c]Change from net return in the base case, [d]Pregnant, live calf: 1% for age 2-12 and 0.4% for age 13 & open to pregnant: 1%, [e]Steer and heifer prices are adjusted.

Table 2.6 Optimal Culling Strategies: Pregnancy Check

Model Results	Net Returns-NR ($)	Change in NR-Base Model	Cull Age		Weighted Average		Head Produced				Replacements Bought	Head Sold	
		%	Open	All	Cow Age	Cull Age	Dam tested as Pregnant	Dam tested as Non-Pregnant	Open tested as Pregnant	Open tested as Non-Pregnant	Bred Heifers	Steers /Heifers	Cull Cows
Pregnancy Check	20,770	68.2	4	13	4.24	6.00	73.1	2.9	3.3	0.0	20.69	44.96	17.3

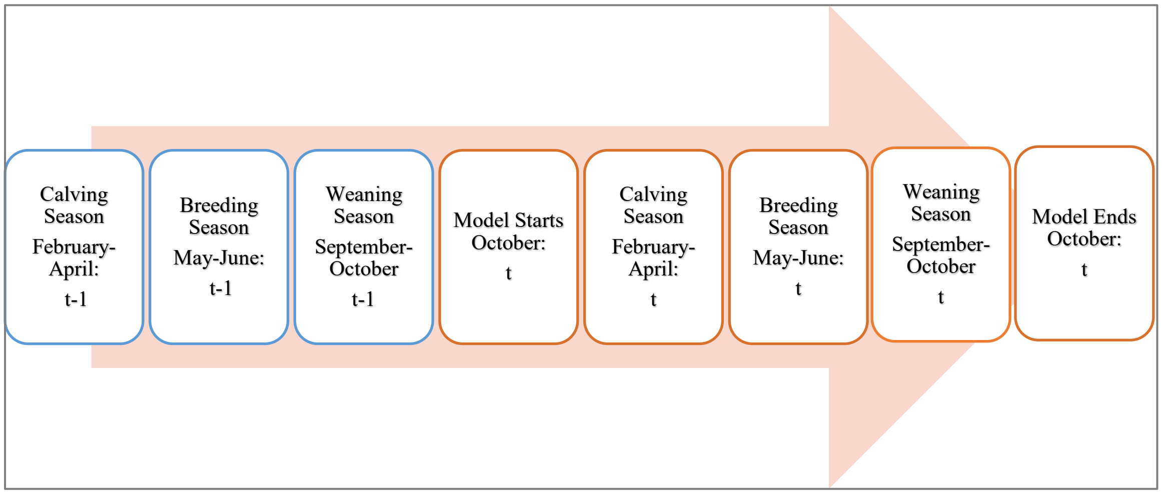

Figure 2.1 Model Timeline (Spring Calving System)

CHAPTER 3. Forecasting Beef-Cattle Prices in the Southern United States: A Model Selection Framework

3.1 Introduction

Price forecasts play an important role in agricultural policy, economic planning, and agri-business management. There is a vast literature using different techniques to forecast agricultural prices (Ates et al., 2019; Darekar & Reddy, 2017; Harris, 2017; Jha & Sinha, 2013; Kurumatani, 2020; Rusiman et al., 2017; Sabu & Kumar, 2020; Wang et al., 2018). Livestock price forecasting literature has had a focus on the distinctive features of the industry such as the biological characteristics of supply, seasonality, cattle cycles, and structural changes. Although time series models including autoregressive integrated moving average (ARIMA) models have been extensively used in the literature, machine learning (ML) techniques have also been examined as an alternative tool to obtain accurate livestock price forecasts. ML techniques have distinct advantages over ARIMA models as they have more flexibility in variable selection and can accommodate both linear and nonlinear behavior in data. In addition to price forecasts, basis forecasts are also common in the literature since they are less volatile and more predictable than spot prices. Basis is the difference between a local cash prices and the relevant futures contract prices for a specific time period. These futures are traded up to a year prior to their expiration date. Since market forces lead to a spot and futures prices, basis is easier to predict with simpler techniques such as moving historical averages (Kastens et al., 1998).

This study aims to develop a model selection framework to forecast beef-cattle prices in Kentucky which is taken as a representative market for the southern region in the United States. The empirical analysis for out-of-sample one-week-ahead forecasts of beef-cattle prices is based on a model comparison and selection process using weekly auction data for steer, heifer, and cull cow cash prices. The contribution of this paper to the current U.S. livestock price forecasting literature is threefold:

1. The study aims to develop a model selection framework to provide beef-cattle spot price forecasts to be used in decision making process by producers. A comparison of traditional time series models (ARIMA, seasonal ARIMA) and ML models is performed to evaluate the difference in the prediction performance of the models in small and large samples and select the best technique to forecast livestock cash prices in the southern U.S.
2. Although there are many studies covering structural and time series modelling in the literature, a limited number of studies use ML models. Given the advancement in statistical learning, ML techniques are useful tools to perform predictive analysis including time series forecasting in the econometrician's toolbox. The study contributes to the current U.S. livestock price forecast literature providing a detailed discussion of selected ML models: ridge regression, LASSO regression, random forests, and gradient boosted machine.
3. This study performs a rolling origin cross validation in model evaluation for both ARIMA and ML techniques. This prevents the data leakage problem by taking the sequential order of time series data into account and avoids arbitrary choice of the test set with a nested algorithm. The rolling origin technique is applied in time series forecasting to ensure that the model does not predict lagged observations using future data during cross validation. The algorithm performed in the paper simultaneously measures prediction error and tunes ML parameters (Cochrane, 2018). This cross-validation strategy provides a useful tool for researchers to obtain

an almost unbiased estimate of the true error in their model estimation process (Varma & Simon, 2006).

Model selection results show that base ARIMA models provide the best forecasts for steer prices and the LASSO regression technique provides better forecasts for heifer and cull cow prices in the southern U.S. The results of this study point to the difference in the prediction power of the ARIMA and ML models in small and large samples and suggest ML techniques as an alternative tool to forecast U.S. livestock prices.

3.2 Literature Review

Prior works in U.S. livestock forecasting have employed structural models, time series models, and ML techniques to find the best forecasting methods and used both cash prices and basis. Maki (1963), an early example from the literature, constructed a multi-equation model which takes the structure of the industry and cattle cycles into consideration. He identified seasonal and cyclical variability as the main source of errors in (two-year and five-year) price forecasts covering 1958-1965. Helmers and Held (1977) used a model comparison approach and they compared the forecasts of eight different models including randomly selecting prices from the previous year, futures prices, past yearly average, linear trend, and U.S. Department of Agriculture (USDA) Outlooks. Although there was not a significant difference between results of methods in terms of error variance, Outlook information provided the best forecasts with the lowest average residual and the smallest standard errors. Bentley and Shumway (1981) estimated a model for adaptive decision making to determine optimal replacement and culling decisions and used price forecasts derived from a cyclical forecasting equation with a cattle price cycle of ten years.

Time series models that use past observations of the same variable for prediction, are mostly preferred as they have a parsimonious approach and control for seasonal and cyclical structure of prices. In their case study for U.S. cattle and hog prices, Harris and Leuthold (1985) conducted a comparative analysis where they evaluate individual and composite techniques including regression (ordinary least squares) and time series models (auto regressive moving average-ARMA and ARIMA). They found that time series models, particularly ARMA, outperform regression models for cattle prices with the lowest root mean square error (RMSE) but do not predict turning points as well as other models. In the presence of nonstationarity, Zapata and Garcia (1990) examined the performance of multivariate and univariate time series with vector autoregressive (VAR) and ARIMA models using monthly U.S. cattle price data. They emphasized the importance of appropriate identification of data under nonstationarity. Their results demonstrated that the ARIMA modelling provides relatively accurate forecasts in the short term and the VAR specification provides more accurate results in the long term. Goodwin (1992) and Guney (2015) contributed forecast methodology with the presence of structural change by using the time varying parameter VAR model to evaluate forecasts for cattle prices. They found that forecast performance improves with a time varying parameter technique. Using model averaging methods, Payne, Karali, and Dorfman (2019) developed a Bayesian model averaging approach to examine cattle basis forecasts at three Georgia locations with weekly data. They pointed out that the Bayesian approach is superior to a set of models including naïve models which takes historical moving averages or the last observed value as the forecast.

Machine learning (ML) techniques including traditional and deep learning models have been used in the literature as alternative techniques for estimation and forecasting. However, ML applications are limited due to data availability and required computational power. Ticlavilca, Feuz, and Mckee (2010) used a Bayesian learning machine approach for regression and artificial neural network (ANN) models for U.S. cattle, hog, and corn prices and stated that Bayesian model outperform ANN models in terms of overall prediction performance and robustness. There are also studies comparing ARIMA and ML methods to determine the best forecasting design for livestock prices. Kohzadi et al. (1996) proposed a model selection framework comparing Neural Network models and ARIMA models to predict U.S. monthly live cattle and wheat cash prices and found that Neural Network models have better performance to forecast livestock cash prices.

This study aims to contribute to the literature with its model selection framework incorporating a rolling origin cross validation scheme to evaluate the predictive performance of ARIMA and ML models and propose the best techniques for steer, heifer, and cull cow prices in the southern U.S.

3.3 Data

Weekly auction prices for Kentucky (KY) and Kansas (KS) from 1993 to 2020 are used in the study and were obtained from USDA Agricultural Marketing Service and Livestock Marketing Information Center. Prices examined are the state averages for 500 to 600 lb medium and large frame #1-2 steers, 450 to 550 lb medium and large frame heifers, and average dressing 80-85% boning cows. Weekly dummy variables are also included to allow for seasonality in the models.

Kentucky raw steer and heifer price samples whose sizes are 531 and 527 respectively, cover the period from 2010 to 2020. As the level of learning in training sets is greatly influenced by sample size, it matters in model selection, particularly in ML applications. Although there is no consensus in the literature for minimum sample size requirement in ML, based on learning curve analysis, prediction accuracy improves and becomes more stable in larger samples (Zhu et al., 2016). On the other hand, the model performance grows logarithmically with sample size for deep learning models (Sun et al., 2017). Furthermore, Cerqueira, Torgo, and Soares (2019) examine the impact of different sample sizes on the performance of traditional time series and ML models. They conclude that traditional time series models outperform ML methods with small samples of less than 100 observations in terms of predictive power, but ML models start to perform better when sample sizes increase beyond 1,000 observations.

As the KY samples are small, a data interpolation process detailed in the Appendix is used to impute missing observations in the samples and increase sample size for KY heifer and steer prices. The purpose is to eliminate sample size bias on model performance, improve performance of ML models, and have a robust model selection procedure. KS steer and heifer prices are used to predict KY prices and VAR causality tests and Structural VAR models are employed to predict statistical relation between the KY and KS series for the period 2010-2020. Then, estimated Structural VAR models are used to generate KY price series for the period from 1999 to 2010. Kansas has the third largest cow-cattle market in U.S. in terms of number of cattle on ranches and feedyards (USDA, 2020). The state has an auction system which is similar to most states including KY in structure and timing. Furthermore, it has more complete datasets than other states. Based on descriptive analysis, KS prices are found to be higher than KY prices for most of the period from 2010 to 2020 for both steer and heifer prices (Figures 3.1a and 3.1b). Apart from the difference in levels, prediction analysis states a meaningful causal relation with a direction from KS prices to KY prices in VAR Granger causality test and SVAR analysis. After these steps, the samples used in steer and heifer forecasting increase to 1,124 observations with KY predicted values. Summary statistics for the data set are presented in Table 3.1.

3.4 Methods

3.4.1 ARIMA Models

ARIMA models with and without seasonality are estimated as baseline standard time-series model comparisons. Seasonality is controlled with weekly dummy variables (Godahewa et al., 2020). Since the number of weeks can vary across years, the 53rd observation is treated as the 52nd when constructing dummy variables (Pan & Yang, 2017). The ARIMA model specified in equation 3.1 assumes a linear relation between KY prices at time t (KY_t), past observations (KY_{t-i}), random errors (μ_t) and moving average component (μ_{t-j}) with a weekly dummy variables matrix (W_k). L represents the lag operator.

$$\left(1-\sum_{i=1}^{p}\beta_i L^{\mathrm{i}}\right)(1-L)^d KY_t = \left(1+\sum_{j=1}^{q}\varphi_j L^j\right)\mu_t + \sum_{k=2}^{52}\gamma_k W_k \qquad (3.1)$$

The ARIMA uses p auto-regressive (AR) terms (β_i), d differencing terms and q moving-average (MA) terms (φ_k) in base ARIMA models. The seasonal ARIMA models contain additional seasonality coefficients (γ_l). The auto.arima() function in the statistical software R is used to estimate ARIMA models with the optimal order for each parameter set using Schwarz information criterion (BIC) to assess model fit and complexity.

3.4.2 Machine Learning Models

A series of supervised learning ML methods are employed in this study. These are divided as a set of shrinkage (ridge and LASSO) and tree-based (gradient boosted machine and random forests) methods. These methods are known as "off the shelf" ML models and commonly used in ML studies. They consider both linear and nonlinear relations among the variables.

3.4.2.1 Shrinkage Methods

Shrinkage methods are a class of linear models that shrink the coefficients toward zero to decrease the variance of predictions (Hastie et al., 2009). Like standard time-series models, these shrinkage models are linear and penalize model complexity but offer more flexible variable selection than standard ARIMAs. The form of shrinkage models, shown in equation 3.2, is similar to regression techniques.

$$KY_t = \beta_0 + \sum_{i=1}^{m} X_{ti}\beta_i + \varepsilon_t \tag{3.2}$$

Here β_0 is the intercept term, β_i is the regression coefficient for $i = 1, \dots, m$, m is dimension of the X_{ti} which is vector of lagged values (KY_{t-i}) and dummy variables (W_l), and ε_t is the error term.

Shrinkage methods penalize model complexity and minimize residual sum of squares (RSS) obtained from equation 3.2 and a penalty parameter which differs by the type of the shrinkage. Ridge regression and LASSO regression are widely used shrinkage techniques in the literature and the train() function in the caret package in R is used with different specifications to estimate ridge and LASSO regression models in the study.

3.4.2.1.1 RIDGE REGRESSION (RIDGE)

The ridge regression uses the sum of squared coefficients as a penalty term and suppresses variable influence by "shrinking" coefficient values. It minimizes penalized RSS in equation 3.3 and allows the model to include all variables.

$$\beta^{ridge} = argmin\begin{Bmatrix} (KY_t - \quad \beta_0 - \sum_{i=1}^{m} X_{ti}\beta_i)^2 + \lambda \sum_{i=1}^{m} \beta_i^2 \\ = RSS + \lambda \sum_{i=1}^{m} \beta_i^2 \end{Bmatrix} \tag{3.3}$$

In the ridge regression, $\lambda \sum_{i=1}^{m} \beta_i^2$ is the shrinkage penalty and $\lambda \geq 0$ is the complexity or tuning parameter. Larger values of λ lead to more shrinkage. The coefficients provide a best fit of the data by minimizing the combination of RSS and the penalty. The trade-off between goodness of fit and model complexity are controlled by λ. When $\lambda = 0$, results correspond with ordinary least squares. As λ approaches infinity, coefficients tend toward zero (James et al., 2017).

3.4.2.1.2 LASSO REGRESSION (LASSO)

LASSO[2] models are similar to ridge regressions in terms of shrinkage method and linearity. Unlike ridge regression, LASSO regressions allow some coefficients of the regression to be equal to zero by using feature selection. The LASSO algorithm minimizes penalized RSS in equation 3.4 and employs a Bayesian variable selection procedure (Storm et al., 2020).

$$\beta^{LASSO} = argmin\begin{Bmatrix} (KY_t - \quad \beta_0 - \sum_{i=1}^{m} X_{ti}\beta_i)^2 + \lambda \sum_{i=1}^{m}|\beta_i| \\ = RSS + \lambda \sum_{i=1}^{m}|\beta_i| \end{Bmatrix} \tag{3.4}$$

[2] LASSO is an acronym for least absolute shrinkage and selection operator.

The minimization strategy in LASSO is similar to the ridge's but the penalty is different. Because the LASSO uses the absolute value of the coefficients in its penalty, its penalty contours have sharp corners at points where each coefficient is zero. This allows LASSO regressions to outright remove variables. Conversely, ridge regressions reduce model variance by reducing the value of the coefficients. However, it does not force coefficients to exactly zero (James et al., 2017).

3.4.2.2 Regression Tree-Based Methods

Regression tree-based methods generally provide better results than linear methods when the linearity assumption is not met in the dataset. They split the variable set into subsets, and then fit a model in each one with different procedures (Hastie et al., 2009). They are nonparametric and provide non-linear predictions, but unlike the shrinkage methods, they do not have an interpretable functional form. This lack of interpretability may limit their usefulness in certain applications where understanding the underlying relationship between variables is important. Random forests and gradient boosted machine methods are used in this study and caret, randomForest, and gbm packages in R are employed to estimate regression three-based models.

3.4.2.2.1 Random Forests (RF)

Proposed by Breiman Leo (2001), the random forest method is robust to small sample sizes, high-dimensional feature spaces, and complex data structures (Tyralis & Papacharalampous, 2017). It is an ensemble method that bases its estimate off of a set of predictions from individual trees. Each tree predicts the outcome variable using a random subset of input variables and a bootstrapped sample of the data. The final predictions of the forest are generated through a process called bagging. Here predictions from individual trees are averaged together. The collective "ensemble" of individual tree predictions increases the stability and accuracy of the model and decreases variance (Divina et al., 2019; Dudek, 2015).

The optimal number of trees can be determined experimentally, the number of input variables in the random selection of subsets, and node size in each tree are parameters to be defined in RF algorithm. The number of input variables in random selection of subsets is suggested as one third of the total variables in the training set by default. It can be changed experimentally if out of bag error decreases. Finally, the number of nodes which determines the depth of the tree is 3 by default and since RF is not too sensitive to this parameter, this number is accepted throughout the literature (Naing & Htike, 2015).

3.4.2.2.2 Gradient Boosted Machine (GBM)

GBM proposed by Friedman (2002; 2001) is another ensemble algorithm which trains a set of regression trees in a sequential procedure and performs an iterative gradient descent algorithm. In the iteration process, the gradient of the loss function is refitted sequentially until no improvement is detected (Divina et al., 2019; Masini et al., 2020). The technique begins with a single decision tree to predict outcome variable based on the input features. The errors of the first tree are then computed by comparing the predicted values and the actual values. The second tree is then constructed to predict the residual of the first tree rather than the original outcome variable. The process is repeated for several iterations, with each new tree aiming to correct the errors of the previous trees. The final prediction value is produced by combining the predictions of all the trees in the ensemble.

The choice of shrinkage parameter, bag fraction, and number of trees are the main practical decisions in the algorithm. The shrinkage parameter determines the learning rate in the procedure. The literature suggests that smaller values (≤0.01) provide better predictions (Friedman 2002). However, the disadvantage of the low values is computational costs in terms of storage and CPU time (Ridgeway, 2020). The bag fraction is the fraction of training set observations that are randomly selected to propose a new tree and is recommended as 0.5 by default (Flores et al., 2017; Ridgeway, 2020). As in random forests, the optimal number of trees in the GBMs can be determined experimentally.

3.4.3 Cross Validation in Time Series Data and Tuning Parameters in ML Models

Cross validation (CV) is a standard process to evaluate model performance to avoid the overfitting problem. Models with almost perfect predictive performance in the training set likely will not perform well with new observations. The CV procedure finds a compromise between model complexity and outside sample fit cross validation error. Proper data splitting into training, validation, and testing sets plays an important role in CV. Storm, Baylis, and Heckelei (2020) suggest that in cases of relatively small sample size as in the case in this study, instead of holding out a separate subsample for validation, k-fold CV can be used as an alternative method. In k-fold CV method, the data set is randomly divided into equal-sized k subsets. The model is fitted with k-1 subsets, validated with the remaining subset, and prediction error is measured with respect to model performance in validation subset. The process is repeated until each of k subsets is used as validation set (Hastie et al., 2009). Determining k in cross validation is important since there is a tradeoff between bias and variance for different numbers of k. Fewer subsets results in higher bias but lower variance. More subsets have lower bias but higher variance and require more computational power. The literature recommends between 5 and 10 validation subsets (Hastie et al., 2009; James et al., 2017).

CV should be applied carefully in time series since there is a risk of data leakage. This occurs when the model uses future observations to predict prior observations and the dependency between observations is ignored (Tashman, 2000). Following Bergmeir, Hyndman, and Koo (2015) and Schnaubelt (2019), a rolling origin cross validation is applied to evaluate a series of ARIMA and ML models. In the rolling origin scheme[3], models are trained with an initial sample size which grows iteratively by shifting forward by one month in each iteration. While the training observations grow across each iteration, the test sample size remains fixed. Therefore, data series are split into only training and test subsamples to train and test models in a CV scheme.

As ARIMA models lack penalty parameters, they are predicted in the CV scheme with only training and test set splits. ML models are trained and validated in the CV scheme with training and evaluation set splits to tune the parameters selected for each ML technique. The training set is split into another training set and a validation set (10% of training set). A nested CV algorithm is used to tune ML parameters (Figure 3.2). The algorithm runs an inner loop for a set of parameter values in the validation set and measures the CV error (sum of square residuals). The parameter value with the minimum CV error is selected as the optimal parameter. The optimal parameter value from the inner loop is then used in corresponding training set to train the model in the outer loop. The process is repeated for each CV fold.

10 cross validation subsets are used in the study. The initial training sample size is 864 observations, and the testing sample size is fixed as 224 observations for each iteration for steer and heifer prices. For cull cow prices, the training sample size is 1,130

[3] It is also known as walk-forward validation.

observations. The test sample size is set as 20% of the data set for each series. Figures 3.3 and 3.4 provide the 10-fold CV scheme for ARIMA and ML models using steer and heifer price series.

The prediction error mean and standard deviation are measured and compared over the CV folds for each model among ARIMA and ML techniques. Prediction error means are the average performance of the technique, and their standard deviations are the variability of the predictive skill of the technique with respect to different samples. At the final step, the technique with the lowest average prediction error measure and standard deviation is selected as the best technique.

3.4.4 Model Selection

RMSE and Mean Absolute Percentage Error (MAPE) shown in equations 3.5 and 3.6 are used to measure the prediction error. Here Y_i and $\hat{Y}_i$ are the real and the predicted values, respectively.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2} \quad (3.5)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right| \quad (3.6)$$

These are commonly used to assess forecasting performance. The models are evaluated with out-of-sample RMSE and MAPE since in-sample accuracy may overstate the model predictive performance (Mullainathan and Spiess 2017).

Lower RMSE and MAPE values mean higher accuracy. RMSE assigns high penalties to large errors, since the prediction errors are squared. This helps to avoid large forecasting errors, but it is sensitive to scale of measurement and data transformation

(Adhikari & Agrawal, 2013). MAPE is computed in percentage terms, normalizes the error by actual values, and is independent of the scale of measurement.

3.5 Estimation Results

3.5.1 ARIMA Models

3.5.1.1 Unit Root Tests

The augmented Dickey Fuller (ADF) and Kwiatkowski, Phillips, Schmidt, and Shin unit (KPSS) tests are applied to check for stationarity and find the level of integration of variables (Dickey & Fuller, 1979; Kwiatkowski et al., 1992). The aim of the KPSS test is to check for stationarity under the presence of a deterministic trend.

According to the ADF test results, the null hypothesis of a unit root is not rejected for all series at any significance level (Table 3.2). All series are determined to be I(1) at the 1% significance level. The KPSS test results reveal that the null hypothesis of stationarity is rejected for all series at levels with both intercept and trend (Table 3.3). The unit root tests using both constant and trend state that all variables are form of I(1). The first differenced series are used in the ARIMA models when the series are I(1).

3.5.1.2 Estimation

Two different ARIMA-type models are estimated: a base model without seasonality and the one with a control for weekly seasonality. The best models for series are selected based on minimum BIC information criteria. Model estimation results are presented in Table 3.4. The base ARIMA technique has lower average RMSE and MAPE values and standard deviations across 10 splits and works better for steer and heifer prices. Adding the seasonality component to the ARIMA design does not improve model performance for steer and heifer prices. However, the ARIMA with seasonality produces better results for cull cow prices, with lower mean predictions errors and lower standard deviations. Peel and Meyer (2002) found that cull cow prices exhibit the largest seasonal patterns among all cattle classes, likely due to liquidation or expansion during the cattle cycle.

3.5.2 Machine Learning Models

3.5.2.1 Tuning the Parameters

In the ridge and LASSO models, the parameter λ is tuned by using a set of 100 λ values. The optimal number of trees for each split is selected in the CV scheme among 500, 1000, and 1500 trees in each model for RF, and among 500, 1000, 1500, 2000, and 2500 trees in each model for GBM. Rf method in R package caret is used to estimate RF models and gbm package in R is used to estimate GBM models.

3.5.2.2 Estimation

The ML algorithms results are presented in Table 3.5. Among ML models, the LASSO algorithm produces models with lower average RMSE and MAPE values for steer, heifer, and cull cow prices.

3.5.3 Model Selection and Discussion

The model selection framework developed in this paper aims to find the best techniques to obtain one-week-ahead beef-cattle cash price forecasts in the southern U.S. Traditional linear ARIMA models with and without seasonality and different ML techniques that allow for linear or non-linear relations in the dataset are compared in a rolling origin cross validation scheme. The impact of sample size on the model performance is also checked by running the techniques with original Kentucky and Kansas samples (Table 3.6 – 3.8).

3.5.3.1 Steer and Heifer Prices

In the larger Kentucky sample, the ARIMA model without seasonality is selected as the best technique to predict steer prices (Tables 3.4 and 3.5). The average prediction errors are 3.77 and 1.69 for RMSE and MAPE, respectively and they are lower than LASSO which is the best ML technique. The LASSO regression prediction error values are 3.91 and 1.86 for RMSE and MAPE, respectively. In the original Kentucky samples with 572 observations, ARIMA without seasonality still has the lowest average prediction errors and outperforms ML techniques (Table 3.6).

In the larger Kentucky sample, LASSO is selected as the best technique to predict heifer prices (Tables 3.4 and 3.5). Its average prediction error values are 3.14 and 1.75 for RMSE and MAPE, respectively and they are lower than both ARIMA-type models. The base ARIMA average prediction error values are 4.35 and 2.60 for RMSE and MAPE, respectively. In the original Kentucky samples with 572 observations, base ARIMA has the lowest average prediction errors and outperforms ML techniques (Table 3.6).

Table 3.7 displays the improvement in prediction performance between best techniques in small and large sample size. As the sample size increases, the difference between the best ARIMA and ML techniques' prediction performance decreases. While base ARIMA model has the best prediction performance in the small sample, the LASSO regression is selected as the best technique to predict heifer prices in the larger Kentucky sample. These results suggest that as sample size increases, ML techniques become preferable over ARIMA-type models (Cerqueira et al., 2019; Sun et al., 2017; Zhu et al., 2016). Model selection results indicate that the base ARIMA models yield the best steer cash price forecasts, and the LASSO regression provides the best heifer cash price forecasts in Kentucky. The difference in the selected best forecasting techniques for steer and heifer prices can be linked to the market structures and mechanisms for price determination. While steers are primarily raised for beef production, heifers are commonly raised for both rebreeding and beef production purposes, which can add complexity to the market structure and mechanism for price determination. Steers typically have a higher feed efficiency than heifers and the cost of gain for steers is lower than for heifers, which can increase their value in the market (Burdine et al., 2014; Martinez et al., 2021; Parish et al., 2018; Williams et al., 2012).

Table 3.8 presents the forecasting results using longer Kansas dataset and the results are consistent when techniques are run with Kansas steer and heifer price samples, consisting of 1,124 observations.

3.5.3.2 Cull Cow Prices

LASSO is selected as the best technique to predict cull cow prices (Tables 3.4 and 3.5). The average prediction errors are 3.79 and 4.33 for RMSE and MAPE, respectively and they are lower than ARIMA-type modes' values. ARIMA with seasonality is the best ARIMA-type model to predict cull cow prices. Its average prediction error values are 4.58 and 5.53 for RMSE and MAPE, respectively. The standard deviations are remarkably low in ML models compared to ARIMA-type models. ML techniques provide lower prediction error variance which results in more precise forecasts with new data. Based on these results, the final best technique is decided as LASSO regression for cull cow prices. The improvement in prediction performance between ARIMA with seasonality and LASSO techniques is substantial, with a 17.25% reduction in RMSE (Table 3.7).

Time series and ML models using dependent variables and lagged endogenous variables provide a parsimonious forecast strategy, particularly in the presence of data availability problems. The high prediction performance of the ML techniques presented in the study makes them useful alternative tools for U.S. livestock price forecasting, especially for heifer and cull cow prices. These ML techniques offer added flexibility in data, providing an advantage over ARIMA models and their added flexibility provides an advantage over ARIMA models (Kohzadi et al., 1996). However, the lack of exogenous variables such as input prices, trade related variables, and interest rates is a potential limitation of the study. Structural models using these variables as explanatory variables may improve forecasting accuracy.

3.6 Conclusion

This study contributes to the livestock price forecasting literature by comparing ARIMA and ML techniques for steer, heifer, and cull cow cash prices in the southern U.S. Different ARIMA and ML techniques are employed using price samples that differ by size and location. To tune, validate, and evaluate models, a rolling origin CV scheme which takes the sequential order of time series into account is used. For ML techniques, a nested cross validation algorithm which performs simultaneous cross validation and parameter tuning is developed and run. Mean and standard deviation of RMSE and MAPE values are compared to decide the best technique. The best models have the lowest average prediction error and standard deviation for each price series and sample.

The results emphasize the importance of the sample size in determining the superiority of ML techniques over standard ARIMA models when forecasting U.S. livestock prices. The model selection results show that ML techniques become preferable over ARIMA-type techniques in larger samples. Model selection results show that base ARIMA models provide the best forecasts for steer prices and the LASSO regression provides better forecasts for heifer and cull cow prices in the southern U.S.

This study's results suggest that ML techniques should be considered as a price forecasting tool for extension specialists and producers to improve their managerial decision-making process. Additionally, the CV scheme proposed in the study offers a strategy for researchers to validate and compare standard ARIMA-type models with newer ML techniques.

Tables and Figures for Chapter 3

Table 3.1 Summary Statistics

Variables	Description	Sample Period	Sample Size	Average ($/cwt)	Standard Deviation
KYS	KY Steer Prices	6/11/1999-12/18/2020	1,124	129.50	44.08
KYH	KY Heifer Prices	6/11/1999-12/18/2020	1,124	120.34	40.22
KYC	KY Cull Cow Prices	2/5/1993-12/19/2020	1,455	52.22	16.77
dumW1 dumW2 …	Weekly dummy variables (e.g., where dumW1=1 in the first week of each year and 0 otherwise)				

Table 3.2 Unit Root Test Results-ADF (t-Statistic)

	Levels Intercept	Levels Intercept and Trend	1st Difference Intercept	1st Difference Intercept and Trend	Result
KYS	-1.99	-2.35	-6.41*	-6.43*	I(1)
KYH	-2.06	-2.28	-7.13*	-7.15*	I(1)
KYC	-2.35	-2.86	-9.467*	-9.463*	I(1)
Critical Values					
1%	-3.44	-3.97	-3.44	-3.97	
5%	-2.86	-3.41	-2.86	-3.41	
10%	-2.57	-3.13	-2.57	-3.13	

Note: *, **, *** indicate rejection of the null hypothesis at 1%, 5% and 10% levels, respectively.
KYS: Kentucky Steer Prices, KYH: Kentucky Heifer Prices, KYC: Kentucky Cull Cow Prices.

Table 3.3 Unit Root Test Results-KPSS (LM Statistic)

	Levels Intercept	Levels Intercept and Trend	1st Difference Intercept	1st Difference Intercept and Trend	Result
KYS	2.54*	0.27*	0.12	0.07	I(1)
KYH	2.29*	0.275*	0.11	0.06	I(1)
KYC	2.16*	0.285*	0.042	0.041	I(1)
Critical Values					
1%	0.74	0.216	0.74	0.216	
5%	0.46	0.146	0.46	0.146	
10%	0.35	0.119	0.35	0.119	

Note: *, **, *** indicate rejection of the null hypothesis at 1%, 5% and 10% levels, respectively.
KYS: Kentucky Steer Prices, KYH: Kentucky Heifer Prices, KYC: Kentucky Cull Cow Prices.

Table 3.4 ARIMA Performance Evaluation*

	Model Selection Criteria		No Seasonality	Seasonality
Steer Prices	RMSE	Mean	**3.77**	3.81
		Standard Deviation	0.45	0.42
	MAPE	Mean	**1.69**	1.71
		Standard Deviation	0.11	0.11
Heifer Prices	RMSE	Mean	**4.35**	4.37
		Standard Deviation	2.69	2.52
	MAPE	Mean	**2.60**	2.63
		Standard Deviation	2.05	1.88
Cull Cow Prices	RMSE	Mean	7.07	**4.58**
		Standard Deviation	3.21	1.38
	MAPE	Mean	8.88	**5.53**
		Standard Deviation	5.00	1.82

*Bold represents the best techniques.

Table 3.5 ML Performance Evaluation*

	Model Selection Criteria		RIDGE	LASSO	RF	GBM
Steer Prices	RMSE	Mean	5.23	**3.91**	6.51	4.57
		Standard Deviation	0.36	0.42	0.30	0.38
	MAPE	Mean	2.70	**1.86**	3.44	2.33
		Standard Deviation	0.14	0.13	0.09	0.13
Heifer Prices	RMSE	Mean	4.67	**3.14**	5.91	3.95
		Standard Deviation	0.25	0.21	0.20	0.29
	MAPE	Mean	2.70	**1.75**	3.65	2.36
		Standard Deviation	0.14	0.08	0.10	0.17
Cull Cow Prices	RMSE	Mean	4.02	**3.79**	5.08	4.63
		Standard Deviation	0.04	0.11	0.15	0.28
	MAPE	Mean	5.08	**4.33**	6.42	5.52
		Standard Deviation	0.12	0.06	0.12	0.15

*Bold represents the best techniques.

Table 3.6 Performance Evaluation: Kentucky Samples with n=572*

	Model Selection Criteria		ARIMA No Seasonality	ARIMA Seasonality	RIDGE	LASSO	RF	GBM
Steer Prices	RMSE	Mean	**3.06**	3.51	4.42	3.83	3.78	3.30
		Standard Deviation	0.04	0.12	0.18	0.05	0.07	0.09
	MAPE	Mean	**1.41**	1.71	2.12	2.05	1.92	1.65
		Standard Deviation	0.02	0.05	0.09	0.07	0.08	0.06
Heifer Prices	RMSE	Mean	**2.92**	3.20	4.11	3.84	4.09	3.20
		Standard Deviation	0.06	0.11	0.20	0.06	0.07	0.10
	MAPE	Mean	**1.51**	1.79	2.33	2.36	2.58	1.87
		Standard Deviation	0.02	0.06	0.12	0.07	0.07	0.10

*Bold represents the best techniques.

Table 3.7 Performance Evaluation: The Improvement in Prediction Performance Between the Best ARIMA and ML Models in Small and Large Samples

		The Best Model	ARIMA	ML	% Reduction in RMSE
Steer Prices	Large Sample-n=1,124	RMSE	**3.77**	3.91	-3.58
	Small Sample-n=572	RMSE	**3.06**	3.30	-7.27
Heifer Prices	Large Sample-n=1,124	RMSE	4.35	**3.14**	-27.82
	Small Sample-n=572	RMSE	**2.92**	3.20	-8.75
Cull Prices	Large Sample-n=1,455	RMSE	4.58	**3.79**	-17.25

*Bold represents the best techniques.

Table 3.8 Performance Evaluation: Kansas Samples with n=1124*

	Model Selection Criteria		ARIMA No Seasonality	ARIMA Seasonality	RIDGE	LASSO	RF	GBM
Steer Prices	RMSE	Mean	**6.46**	6.65	6.82	6.58	8.16	6.73
		Standard Deviation	0.04	0.11	0.06	0.15	0.17	0.32
	MAPE	Mean	**2.89**	3.0348	3.139	3.0351	3.95	3.143
		Standard Deviation	0.04	0.09	0.03	0.06	0.12	0.15
Heifer Prices	RMSE	Mean	5.72	5.84	6.29	**5.40**	7.38	5.79
		Standard Deviation	0.10	0.14	0.06	0.07	0.22	0.24
	MAPE	Mean	2.94	3.02	3.30	**2.77**	3.97	2.97
		Standard Deviation	0.09	0.09	0.03	0.03	0.16	0.14

*Bold represents the best techniques.

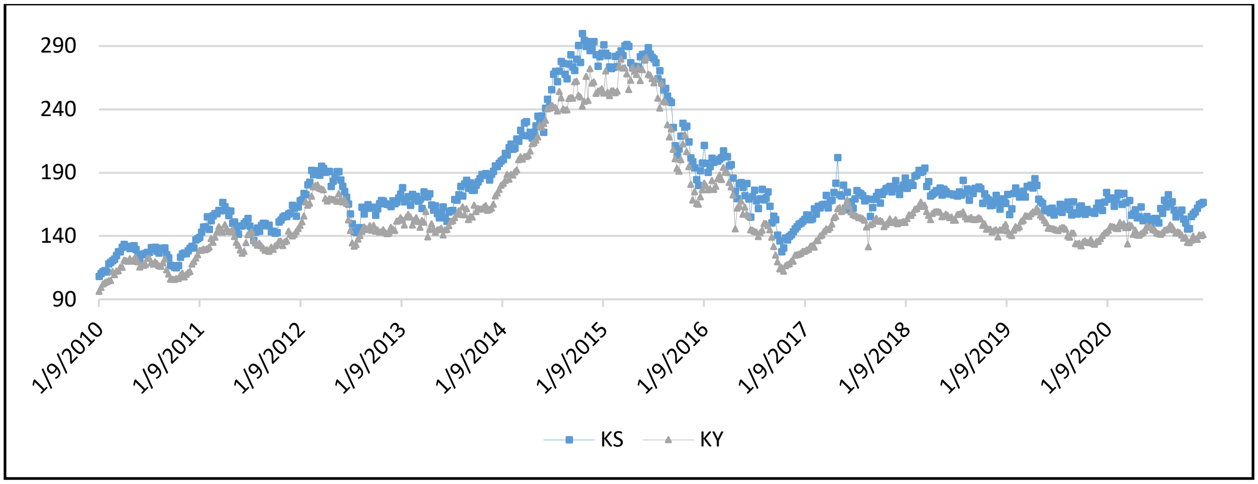

Figure 3.1a KY and KS Steer Prices

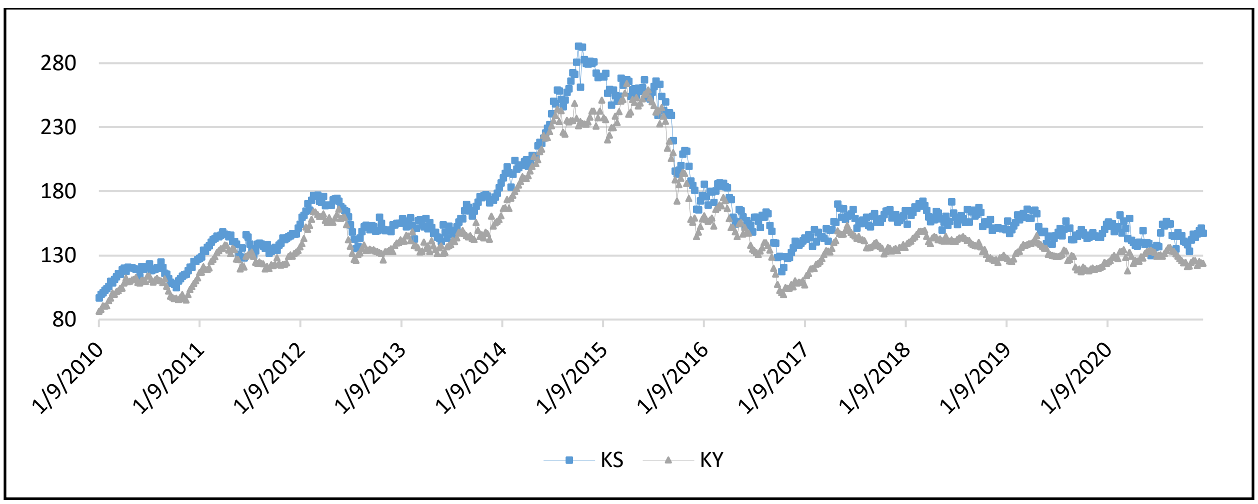

Figure 3.1b KY and KS Heifer Prices

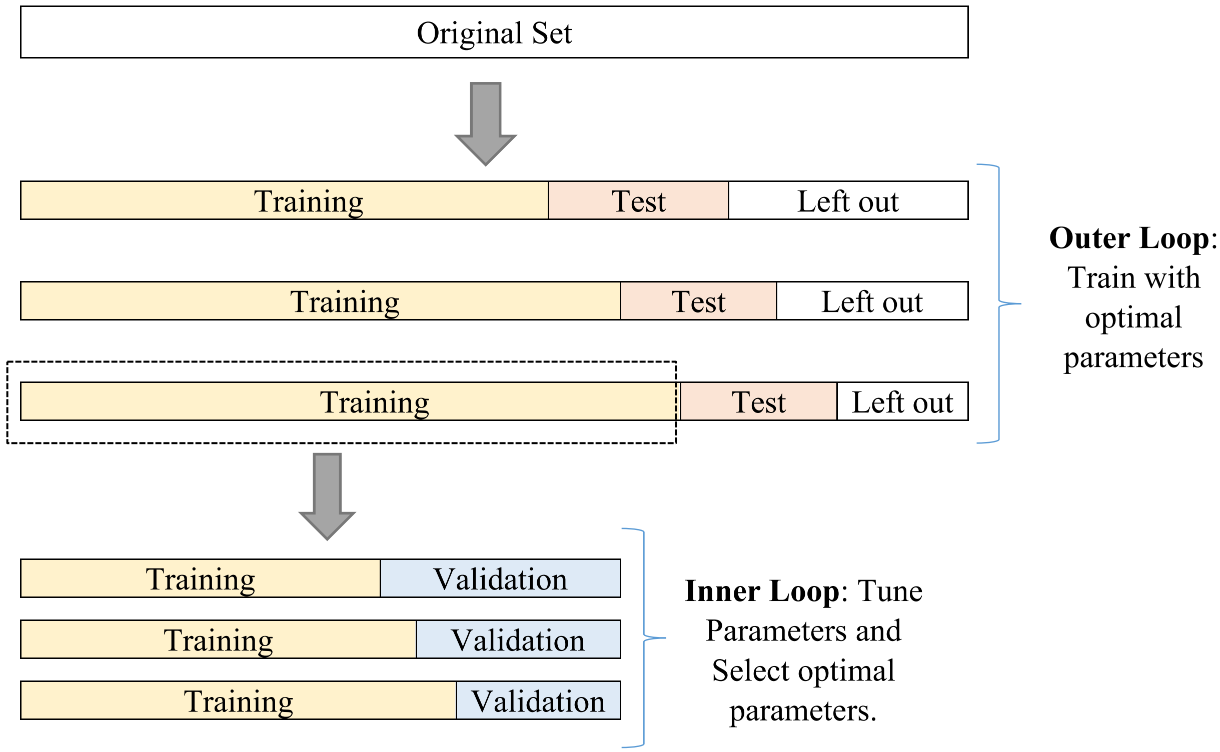

Figure 3.2 Nested Cross Validation Algorithm for ML Models

Initial Training Sample: **864** Fixed Test Sample: **224**, Shift: **4** weeks, Iteration: **10**			
Iterations	Training	Test	Left out
1	**864**	**224**	**36**
2	868	224	32
3	872	224	28
4	876	224	24
5	880	224	20
6	884	224	16
7	888	224	12
8	892	224	8
9	896	224	4
10	900	224	
	Number of Observations		

Figure 3.3 Cross Validation Scheme: ARIMA Models

Initial Training Sample: **864**, Fixed Validation Sample: **84**, Fixed Test Sample: **224**, Shift: **4** weeks, Iteration: **10**				
Iterations	Training	Validation	Test	Left out
1	**780**	**84**	**224**	**36**
2	784	84	224	32
3	788	84	224	28
4	792	84	224	24
5	796	84	224	20
6	800	84	224	16
7	804	84	224	12
8	808	84	224	8
9	812	84	224	4
10	816	84	224	
	Number of Observations			

Figure 3.4 Cross Validation Scheme: ML Models

CHAPTER 4. THE COVID-19 SHOCK AND DYNAMICS OF PRICE ADJUSTMENT IN THE U.S. BEEF SECTOR[4]

4.1 Introduction

SARS-CoV-2 which turned into coronavirus disease in 2019 (COVID-19), was first detected in December of 2019 in China. The global impact was realized in March 2020 after the World Health Organization (WHO) announced it as a global pandemic. COVID-19 has caused devastating health, social, and economic problems across the world, unlike other members of coronaviruses family: severe acute respiratory syndrome-SARS and Middle East respiratory syndrome-MERS (National Institute of Allergy and Infectious Disease, 2021). The first federal policy response in the U.S. was the declaration of a national emergency by President Trump on March 19, 2020. On March 27, 2020, the U.S. Senate passed a $2 trillion Coronavirus Aid Relief, and an Economic Security (CARES) Act to support hospitals, small businesses, and state and local governments.

Supply chain disruptions during the COVID-19 pandemic led to unexpected price movements in the agricultural markets. The supply chain disruptions in the U.S. beef market were first experienced at the beef packing plants and processors due to the spread of COVID-19 among the workers. This caused temporarily shutdowns of some plants and capacity limitations at some facilities (Balagtas & Cooper, 2021; Bunge, 2020). These shutdowns and capacity shortages at the wholesale level resulted in price hikes at the wholesale and retail levels, and oversupply, lower prices, and income losses at the farm level (Cowley, 2020).

[4] This chapter is reproduced from Erol and Saghaian (2022).

Figures 4.1 and 4.2 depict the changes in the U.S. beef prices and price spreads. Price hikes were observed between March and May 2020 for wholesale prices and between March and June 2020 for retail prices (Figure 4.1). The increase was 58.9% in wholesale prices and 17.7% in retail prices in May 2020. Although price decreases at farm level started before pandemic, it worsened between March-July. The impact on price spreads is also worth mentioning. Intra-chain price spreads are the difference between prices at different stages and useful tools to measure efficiency and equity of the beef supply chain (Pouliot & Shulz, 2016). Wholesale-farm price spread started at 98 cents in May 2020, and jumped to 389 cents. Although it decreased to 97 cents after two months, it remained above the average value of the period January 2016-April 2021 (Figure 4.2). The opposite movements in the retail-wholesale price spreads were observed. Retail-farm and retail-wholesale price spreads stayed above their averages after the initial pandemic effect.

In this study, the type, magnitude, and speeds of price adjustments of the COVID-19 shock along the beef marketing channel are investigated with different econometric methods. A vector error correction (VEC) model with structural breaks and historical decomposition graphs is used to investigate the impact of the COVID-19 pandemic shock on the short-run dynamics and speeds of price adjustment along the U.S. beef marketing channel to examine how the pandemic affected the adjustment patterns along the beef supply channel. In addition, the long-run relationships among the farm, wholesale, and retail beef prices are analyzed. The long-run convergence of the price series could point to market efficiency and integration of the beef marketing channel. Price adjustment along the farm, wholesale, and retail levels, and its impact on the economic agents across the beef marketing channel determine market structure and market efficiency of the U.S. beef market. The historical decomposition graphs provide the dynamic effects of the impact of the COVID-19 pandemic on the price series in a neighborhood of the event.

Although there is an extensive literature on price transmission in the U.S. beef markets, the number of the studies focusing on the impact of the COVID-19 pandemic on the adjustment processes is limited. Since the process is dynamic, analyzing the impact of external shocks such as COVID-19 would also contribute to the understanding of price discovery in the U.S. beef market. An understanding of price relations along the U.S. beef supply chain provides vital information for the beef market structure, producers' welfare as well as public policy.

The results of this study indicate asymmetric price adjustment in the U.S. beef supply channels, both in speed and magnitude. The empirical results show that U.S. beef markets return to their pre-shock patterns in 4 to 6 months. The results also reveal that the impact of the COVID-19 shock is uneven across the beef supply chain with farmers bearing most of the burden of the shock. The historical decomposition graphs demonstrate retailers and wholesalers having higher prices, while farmers receiving lower prices than their predicted values during the COVID-19 pandemic.

4.2 Literature Review

There is an extensive literature on the impact of the COVID-19 shock and policy responses in the agriculture sector, the U.S. and world economies (e.g., Baldwin & di Mauro, 2020; Bairoliya & İmrohoroğlu, 2020; Laborde et al., 2021; Beckman & Countryman, 2021). The state of agricultural food markets and supply chains under the COVID-19 pandemic are evaluated mostly with respect to demand shocks and supply channel disruptions. The immediate impacts are mostly tied to price changes, hoarding and changes in the consumer behavior, decreases in incomes and job losses, food shortages, and increased uncertainty and recession risks under the pandemic conditions (Poudel et al., 2020; Swinnen & Vos, 2021). Although these impacts differed across countries and regions, lockdown policies, labor shortages, farm-level productivity reductions, and international trade restrictions under COVID-19 disrupted supply chains across countries and regions (Barman et al., 2021; Laborde et al., 2021; Swinnen & Vos, 2021; Van Hoyweghen et al., 2021).

Siche (2020) stated that high-value products like meat and perishables have had more significant price spikes. Mead et al. (2020) evaluated price movements during the March-June 2020 period and confirmed that meat products had large price increases during the pandemic. Balagtas and Cooper (2021) discussed the impacts on the U.S. meat markets in terms of both domestic and international dynamics and pointed out that global trade restrictions put a downward pressure on the prices, and private precautions with mandated shutdowns altered demand structure in favor of food at home.

There are also many vertical price transmission studies in the literature. The primary goal in the vertical price transmission literature is to identify the response of market participants at different stages of the supply chain to price and policy changes. Whether the shocks are passed asymmetrically with unequal magnitude and speed throughout the supply chain tells a great amount about the efficiency and equity of the beef supply chain. In an efficient market, price transmission is complete, and the value created in the supply chain is distributed equally among market participants. To explain these issues, these studies have applied different econometric techniques with different data frequencies for upstream and downstream stages of supply chains of various products for different countries. Meyer and Von Cramon-Taubadel (2004), Vavra and Goodwin (2005), and Lloyd (2017) provide a detailed review on the literature in agri-food markets.

The results of these studies depend on the models and data used in these studies, and the effects of COVID-19 in the U.S. beef markets remain unclear with mixed conclusions. The assumptions related to the type of adjustment i.e., whether it is symmetric or asymmetric, direction of price transmission, long run model specification, possible structural breaks, impact of exogenous shocks, and varying data types are the main differences among these studies. Most of these studies have investigated asymmetries with linear and nonlinear models such as VEC and threshold vector error correction (TVEC), threshold autoregressive (TAR), momentum threshold autoregressive (MTAR), and nonlinear autoregressive distributed lag (NARDL) models.

Using TVEC model with weekly data for the period January 1981-March 1998, Goodwin and Holt (1999) found a unidirectional causal relation from farm to wholesale to retail level and concluded that adjustment was symmetric, which implies an efficient market structure in the U.S. beef market. Paying a special attention to the data type, Rojas, Andino, and Purcell (2008) used Bureau of Labor Statistics (BLS) and scanner data for retail level prices in a unidirectional relation in a VEC model. Based on the model coefficients and impulse response functions, they found a symmetric adjustment between retail and wholesale prices for both BLS and scanner data and stated that there is a quicker adjustment with larger magnitude in scanner prices to wholesale price changes comparing to BLS data. The relatively short sample size of 56 observations and a questionable representation power of their scanner data are counted as two drawbacks of the study.

Boetel and Liu (2010) contributed to the literature with structural breaks in a long run cointegration relation among the price series. They investigated possible structural breaks in a unidirectional long-run relation from farm to wholesale to retail level for beef and pork prices endogenously and used these structural breaks in a form of structural dummies in an asymmetric VEC model to test for asymmetry in price transmission in monthly price data from January 1970 to February 2008. Their results indicated an asymmetrical speed of adjustment and a bidirectional relation for the price series. Surathkal et al. (2014) employed TAR and MTAR models for monthly wholesale and retail prices and augmented their models for product cuts and quality grade differences. Their results showed significant asymmetries with different effects of decrease or increase in the wholesale prices on retail prices and confirmed the variation with quality grades.

Emmanouilides and Fousekis (2015) applied a copula-based modelling to test the degree of price dependency between monthly farm and wholesale, and wholesale and retail price data for the period from January 2000 to June 2013. They detected strong positive asymmetric price transmission between farm and wholesale prices and pointed out market power and efficiency concerns in the U.S. beef markets. Fousekis, Katrakilidis, and Trachanas (2016) used a NARDL model to farm-wholesale and wholesale-retail prices with monthly data from January 1990 to January 2014. They concluded the existence of asymmetry in magnitude for the farm-wholesale price transmission and the presence of asymmetry both in speed and magnitude for the wholesale-retail price transmission. Their results implied an advantage for wholesalers over farmers and for retailers over wholesalers.

Pozo, Bachmeier, and Schroeder (2021) focused on the importance of the data type in the price transmission studies, using BLS and scanner datasets with both weekly and monthly frequencies for retail level prices. The authors used a TVEC model with a unidirectional relation from farm to wholesale to retail levels. They concluded that symmetry existed in scanner data, implying efficiency of U.S. beef markets, but detected asymmetry in the BLS data.

Although there is a vast literature testing price transmission asymmetry in the vertical chain of U.S. beef markets, number of studies which investigate these dynamics under external shocks is limited. Livanis and Moss (2005) and Saghaian (2007) investigated the impact of food safety scares on price spreads and adjustments in the U.S. beef markets. Using monthly farm, wholesale, and retail price data, Livanis and Moss (2005) employed impulse response functions analysis based on a symmetric VEC model specification with structural break dummies for farm and wholesale prices and an impulse dummy for Food Safety Index. They used the index as a proxy variable for the impact of mad cow disease on consumer behavior in beef purchasing. Their analysis concluded that each of these prices has a different response to a shock from the food safety index. They stated that retail prices are less responsive, while farm and wholesale prices are more responsive with a longer recovery period.

Saghaian (2007) investigated the case of Bovine Spongiform Encephalopathy (BSE) discovery in the U.S. beef sector and applied symmetric VEC model with weekly farm, wholesale, and retail price series. The study found a bidirectional transmission with asymmetric adjustment in both speed and magnitude among different stages. The study also performed historical decomposition analysis to measure the impact of the shock. The results manifested a differential impact of the shock, which increased price margins, particularly at the retail price level. The outcomes of the study emphasized concerns about the efficiency of the U.S. beef markets.

Darbandi and Saghaian (2016) estimated the impact of Great Recession with a symmetric VEC model and historical decomposition analysis for monthly price series for three stages of the beef supply chain. They found evidence of asymmetric price adjustment in both speed and magnitude, indicating the inefficiency of the beef supply chain. Furthermore, their historical decomposition results demonstrated a positive impact of the Great Recession with significant difference across the stages.

Ramsey et al. (2021) estimated both linear and threshold models with weekly data ending in July 2020 for pair of wholesale-retail price transmission, and analyzed the dynamic relations in chicken, beef, and pork markets along with the impact of COVID-19, using an event study. They concluded that while retail and wholesale prices have different speeds of adjustment, the immediate shock of COVID-19 is transitory and U.S. beef markets are well functioning, with prices returning to their predicted levels quickly.

In this study, the dynamic price relations in the U.S. beef supply chain are examined in a symmetric VEC model and a historical decomposition analysis is applied to measure the impact of COVID-19 pandemic on the price adjustment process. The study's contribution to the current literature is twofold: the methodology used accounts for endogenous structural breaks in the long-run cointegration relations of price series for the period from January 1970 to April 2021, and it estimates the impact of COVID-19 on all the stages of the beef supply chain, including the recent periods.

4.3 Methods and Data

The empirical VEC model is specified in equation 4.1:

$$\Delta P_t = \alpha_0 + \sum_{i=1}^{k-1} \Gamma_i \Delta P_{t-1} + \Pi P_{t-1} + \sum_{j=2}^{12} \gamma_j M_j + \sum_{l=1}^{4} \lambda_l D_l + \varepsilon_t \qquad (4.1)$$

Where P_t is a p-element vector of observations on three endogenous price variables in the system at time t, α_0 is a vector of intercept terms, $\Gamma_i \Delta P_{t-1}$ term accounts for the short-run relationships among the price series, and Π matrix contains the long-run cointegration relationship. P is a 3×1 vector since there are three price series. M_j is seasonal dummy variables which are used to gauge more accurate pattern in predictions in historical decompositions and D_l is structural dummy variables. ε_t is the error term with zero mean and non-diagonal covariance matrix.

VEC is a vector autoregressive (VAR) model in the first difference form that is suitable to estimate the relationship between non stationary-I(1) series which are stationary-I(0) after first differenced if their linear combination is I(0) (Engle & Granger, 1987). That means they are cointegrated and deviations from equilibrium are stationary. These conditions allow us to use an error correction specification to model their relation. The Johansen (1991) procedure is applied to test for the existence of cointegration and model estimation.

Herein, the traditional Augmented Dickey Fuller (ADF) and the Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) tests are employed to check the stationarity of series. The rejection of the null hypothesis in ADF test means that the series is stationary, and the mean and variance are stable over time (Dickey & Fuller, 1979). The KPSS test, which has a null hypothesis of stationary is performed since the inclusion of a trend term may reduce the power of the ADF test (Kwiatkowski et al., 1992; Özertan et al., 2014). As these tests do not account for structural breaks in time series, and results may fail due to structural breaks in the series, the Zivot-Andrews test which allows for endogenous structural breaks is used to avoid the impact of structural breaks on unit root tests (Zivot & Andrews, 1992). The rejection of the null hypothesis means that the series is stationary under a structural break.

To have a proper model specification in the VEC model under the existence of structural breaks, structural breaks are checked in the long run equilibrium relationship specified in equation 4.2 which follows a direction from farm prices to wholesale and retail prices based on pairwise causality tests.

$$R_t = \beta_0 + \beta_1 * W_t + \beta_2 * F_t + v_t \tag{4.2}$$

Here R_t, W_t, and F_t denotes retail, wholesale, and farm prices, respectively. Bai (1997) is used to detect multiple unknown structural breaks in the long run equation. The process is as follows; the algorithm starts from the whole sample and performs a test which has a null hypothesis of constant parameters. If the null hypothesis is rejected, then it divides the sample into two subsamples at a break point. The algorithm applies the test to both subsamples, and it estimates another break in case of rejection of the null hypothesis. The process ends when subsamples do not reject the null hypothesis.

Structural breaks are modelled in VEC specification with different methods in the literature. After detecting structural points in the series with Zivot-Andrews and Clemente-Monates-Reyes unit root tests with thresholds, Pala (2013) divides the full sample into two subsamples at the break points and estimate two separate VEC models to account for the impact of the structural breaks on the cointegration relationship between crude oil and food prices. Özertan, Saghaian, and Tekgüç (2014) and Livanis and Moss (2005) construct structural dummies for individual series and specify their models with dummy variables. Following Boetel and Liu (2010), structural dummies for the break points are detected in the long run equation and the VEC system is specified accordingly.

Finally, historical decomposition graphs are used to measure the impact of the COVID-19 pandemic. The historical decomposition function tracks the evolution of beef prices through the system and breaks down the price series into historical shocks in each series to determine their responses in a neighborhood (time interval) of the event (Chopra & Bessler, 2005). This method specified in equation 4.3, decomposes each price series to determine the impact of the shock on prices in a neighborhood of the event:

$$\mathrm{P}_{\mathrm{t+j}} = \sum_{\mathrm{s=0}}^{\mathrm{j-1}} \psi_s U_{t+j-s} + \left[X_{t+j}\beta + \sum_{\mathrm{s}=j}^{\infty} \psi_s U_{t+j-s}\right] \tag{4.3}$$

Where $\mathrm{P}_{\mathrm{t+j}}$ is a multivariate stochastic process, U is its multivariate noise process, X is the deterministic part of $\mathrm{P}_{\mathrm{t+j}}$, and s is a counter for the number of time periods (Regression Analysis of Time Series, 2010). The first part of equation 4.3 represents the part of $\mathrm{P}_{\mathrm{t+j}}$ that is due to the shock, and the second part is the forecast of price series based on the information available at time t, the date of the event (Saghaian, 2007).

In historical decomposition graphs, each series in the representative VEC model are partitioned into two parts: one is due to innovations that drive the joint behavior of beef prices for period t_1 to t_j, the horizon of interest, and the other is the forecast of price series based on information available at time t, the date of the COVID-19 pandemic event. This latter forecast of prices reflects how prices would have evolved if there had been no COVID-19 shock. It traces the response of forecasted prices to the beef price innovations in the absence of a shock as well as actual values in the presence of the shock. Hence, the historical decomposition equation estimates the percentage deviation in the actual prices explained by the shocks compared with the forecasted prices. A potential limitation of this method is that the effect of other shocks on beef prices during the COVID-19 pandemic is ignored or assumed to be accounted for in the estimated models. Due to the substantial impacts of the pandemic given in the study and the period selected in historical decomposition analysis, the method is still a useful tool which enables evaluation of the impact of the pandemic on the markets and provides insights to the dynamics of markets during the pandemic.

The dataset used in this study covers monthly farm, wholesale, and retail price data from January 1970 to April 2021. The price data is obtained from USDA, Economic Research Service. Table 4.1 provides summary statistics for both level and logarithm form of data. All prices are related to Choice grade and in cents per pound. Economic Research Service calculates retail value as a weighted average of retail meat cuts and wholesale value as the value of the meat when it leaves the packing plants. The farm value represents the value of the meat to the farmers, and it is equal to the difference between the value of the cattle and by products. Natural logarithm of prices is used in empirical analysis.

4.4 Estimation Results and Discussion

4.4.1 Unit Root Tests

Unit root test results are presented in tables 4.2 and 4.3. Two specifications of the trend function are used in all tests. One includes only the intercept term, while the other one has both intercept and trend terms. The results show that retail prices are stationary after taking the first difference. Although ADF results state that wholesale prices are stationary with trend, KPSS test concludes that the series is stationary after first difference.

The KPSS test concludes that farm prices are stationary with trend. The results of the Zivot-Andrews test show that all series are stationary at the first difference. All series have break points in June 1993, when tested at levels with only the intercept term. The test results confirm the existence of structural breaks in the individual series and suggest that model specifications incorporating structural breaks would yield better results compared to models that do not account for the breaks.

4.4.2 Structural Breaks in the Long-run Cointegration

The results of pairwise Granger causality tests confirm a unidirectional price transmission in the U.S. beef supply chain and the direction of causality is from upstream to downstream (Table 4.4). These results state that retail markets are responsive to the shocks at farm and wholesale levels.

The structural break test applied to the long run equilibrium relationship specified in equation 4.2, detects four structural breaks (Table 4.5). Heteroscedasticity and autocorrelation consistent covariance matrix tests were used in the estimation of the equation for a maximum of 5 breaks with a 15% trimming rate. The null hypothesis in the test checks the significance of l break points against l+1 break points, and critical values are obtained from Bai and Perron (2003). As shown in Table 4.5, the four estimated structural break dates are November 1980, July 1993, May 2001, and September 2013.

Boetel and Liu (2010) also detected similar dates with lagged specification in the long-run equation. They used different datasets for farm level prices, and lagged variable of farm prices in their specification and their period ended in February 2008; hence the year of breaks are similar. They argued that 1978 energy prices in 1981 break, the trade regime changes in the U.S. exports in 1993, and the impact of Atkins diet phenomenon on the beef industry in 2001 could be the reasons for the breaks. The study herein similarly relies on their reasoning for the break points. The break in 2013 might be related to the herd contraction that started in 2010 and worsened all through 2013 because of unfavorable dry conditions. These break dates divide the sample into 5 different regimes. Structural dummies are constructed for these regimes in the VEC model specification. Each dummy takes value of 1 from its starting period until the beginning of the other break, and 0 for the other periods. For the model in this study, dummies are specified only to allow intercept shifts because price series are nonstationary.

4.4.3 The Johansen Cointegration Test and VEC Model Estimation

This study follows the Johansen (1991) testing procedures to specify a cointegration model including intercept and slope coefficient consistent with the underlying data generation process. Test results are provided in Table 4.6. At the 5%level of significance for the trace test, the null hypotheses that rank is equal to 0 and 1 is rejected. However, there is a failure to reject the null hypothesis that the cointegrating rank of the system is at most two at the 5% level. These results confirm that there are two long-run equilibrium relationships between the series. The cointegration relation assures that there is a long-run relationship among the series. Hence, the model can empirically address the recovery of the deviation from long run equilibrium with the speed of adjustment.

The optimal lag length for the VEC model is selected as one based on the underlying VAR model and Bayesian information criteria. The Durbin-Watson bounds test confirms that there is no evidence of the first-order autocorrelation at the 5% level of significance. The stability of the model is also checked with characteristic roots and ensure that they have modulus less than one and lie inside the unit circle. Model results are provided in Table 4.7. The R^2 values indicate that between 17% and 38% of the variation in the price series are explained by the models. The speeds of adjustment for all series are statistically significant.

The speed of adjustment for the wholesale prices is much higher (0.21), in absolute value, than both farm prices (0.074) and retail prices (0.021). This is an indication of asymmetric price transmission with respect to speed among the different stages of the U.S. beef supply chain. The result shows that following the COVID-19 shock, wholesale prices adjusted more quickly than both farm (threefold) and retail prices (tenfold). This suggests that wholesale prices were more flexible than retail and farm prices to restore to the long run equilibrium with the COVID-19 shock. This result is consistent with the results of Saghaian (2007) and Darbandi and Saghaian (2016), who used a similar methodology. However, it rejects the conclusion of Goodwin and Holt (1999) whose results showed a symmetric adjustment. Ramsey et al. (2021) also found an asymmetric relation between retail and wholesale prices, but the speed of adjustment for retail prices was larger, having a higher speed of adjustment compared to the wholesale prices using weekly data.

The economic literature accounts for a variety of reasons for the asymmetric adjustment in the U.S. beef markets. Balagtas and Cooper (2021) discussed the market power exercised by meatpackers during the COVID-19 pandemic and concluded that meatpackers took advantage of their market power to increase price margins. In addition to the market power, product heterogeneity, long-term contracts, and adjustments or menu costs are other reasons stated for the existence of asymmetric price adjustment along the U.S. beef supply chain (Goodwin & Holt, 1999; Saghaian, 2007; Zachariasse & Bunte, 2003).

4.4.4 Historical Decomposition Graphs

The empirical results show that there is asymmetric adjustment in speeds of prices in the U.S. beef supply chain. To measure the magnitude of price adjustment, historical decomposition graphs are employed. These graphs show the short-run dynamic effects of the COVID-19 shock on the prices in the neighborhood of the event, i.e., the COVID-19 shock. Figures 4.3a-c provide the historical decomposition graphs of the price series for 15 months of the forecast horizon. Herein, it is assumed that the initial impact of COVID-19 on the prices started in March 2020, when the official policy responses and decisions took stage. Before April 2020, the actual and predicted data show almost the same patterns. However, significant differences between actual and predicted prices start to emerge after that period.

The historical decomposition graphs of predicted prices show positive changes for the retail and wholesale prices, but a negative impact for the farm level prices. Actual retail and wholesale prices rose above their predicted values and actual farm prices fell below their predicted values following COVID-19 shocks. The historical decomposition graph of the wholesale prices, that includes the impact of the shock, presents a wide departure of actual prices occurring immediately in April 2020 and continuing until July 2020. The maximum deviation from the predicted value is almost 10% in May 2020. The wholesale prices approach their predicted levels after 4 months.

The historical decomposition graph of the retail prices also shows a positive change with divergence of actual prices starting in April 2020 and continuing until August 2020. The maximum deviation from predicted value is almost 3% in May 2020. The adjustment process of actual prices takes about 6 months to converge to the predicted price values.

However, actual prices stay slightly above the predicted values after 6 months. These results, consistent with the results for speeds of adjustment, indicate a differential impact of COVID-19 on wholesalers and retailers.

More severe differential impact is observed when analyzing the historical decomposition graph for the farm prices. As mentioned earlier, farm prices likely fell due to the COVID-19 shock. The negative divergence of actual prices starts in April 2020 and ends by July 2020. The maximum deviation from predicted value is almost -4% in July 2020. These results imply that in the short run, an exogenous COVID-19 shock on the U.S. beef sector impacted cattle producers negatively, while it positively affected packers and retailers. Although the initial time of the shock is similar at all stages, its impact was felt at different periods across the series. The magnitudes of price effects are substantially different for the price series, resulting in widening the producer-retail price spreads.

The results confirm an asymmetry in the magnitude in the U.S. beef prices. The outcomes of the event study in Ramsey et al. (2021) also find similar results for retail and wholesale prices and they cite the supply shortage during the pandemic as a reason for price spikes during April and May 2020 as stated early. They define these shocks as transitory and point out that the transition period is 1 to 3 months; results of the study herein found 4 to 6 months. Although some minor differences between actual and predicted prices that imply price stickiness and incomplete price transmission after the shock period for some of the price levels can be experienced, it can be said that U.S. beef markets are resilient enough to absorb the shocks and return to their pre-shock patterns in 4 to 6 months. Another point is that farmers are the only and most adversely impacted economic actors in the U.S. beef supply chain during the COVID-19 shock. This outcome rationalizes the base for policy makers to prioritize farmers in support policies during similar crises such as COVID-19.

4.5 Conclusion

The U.S. beef industry has drawn a considerable amount of attention from research institutions and academia. Apart from its size and economic impact, the structural changes in the production and market efficiency, alongside the federal and state level supports, add to the importance of the price discovery and price adjustments in the beef sector.

This study proposes a methodology to investigate how COVID-19 shock impacted the U.S. beef supply chain and examines whether dynamic price relations among retailer, wholesaler, and farmers changed during this historical shock. Contemporary time-series techniques are employed in a VEC methodology augmented with structural breaks to measure the speeds of price adjustments, and historical decomposition graphs to estimate the magnitude of the adjustments across the supply chain with monthly prices for the period from January 1970 to April 2021. The assumed actual shock period started in March 2020 after the spread of COVID-19 and initiation of related policy reactions in the U.S.

The empirical results provide several implications to contribute to the current literature and construct a base for strategic agribusiness reactions for similar crises. First, in this study, the capability of markets to recover after a shock like COVID-19 is evaluated. Results suggest that the U.S. beef markets are resilient enough to absorb the shocks and return to their pre-shock patterns in 4 to 6 months. To obtain more accurate results, structural breaks are incorporated in the model estimating the dynamic price relations in the U.S. beef markets. Model results show that the direction of price relationships is from farmers to wholesalers, to retailers in this study's dataset. Furthermore, the VEC model results indicate that there is an asymmetry in the speeds of price adjustments during the COVID-19 shock. The wholesale prices adjusted more quickly than farm (threefold) and retail prices (tenfold). That suggests higher flexibility of wholesale prices compared to the retail and farm prices when prices restore to the long run equilibrium. These results have welfare and policy implications for the U.S. beef industry.

In addition, historical decomposition graphs show the existence of asymmetry in the magnitude of the price adjustments. This is consistent with the differential speeds of adjustments discovered. The impact causes retail and wholesale prices to be higher than their predicted values. That is, price spreads are widened due to the COVID-19 shock in favor of wholesalers and retailers. Hence, the shock has adversely affected the consumers in the U.S. beef marketing chain.

Meanwhile farm prices are lower than their predicted values due to the COVID-19 shock. That has an important agribusiness implication for the farmers in the U.S. beef supply chain. This study concludes that beef producers' incomes are highly adversely impacted because of the shock and should be prioritized by policy makers during similar crises like COVID-19.

Tables and Figures for Chapter 4

Table 4.1 Summary Statistics

Level	Retail	Wholesale	Farm
Mean	330.00	198.48	166.95
Median	283.85	176.00	150.85
Maximum	758.51	638.56	367.02
Minimum	98.00	71.50	58.80
Std. Dev.	152.98	78.10	63.25
Observations	616	616	616
Log	Retail	Wholesale	Farm
Mean	5.69	5.22	5.05
Median	5.65	5.17	5.02
Maximum	6.63	6.46	5.91
Minimum	4.58	4.26	4.07
Std. Dev.	0.49	0.39	0.38
Observations	616	616	616

Table 4.2 ADF and KPSS Test Results

ADF t-statistic (BIC lag)	Levels Intercept	Levels Intercept and Trend	1st Difference Intercept	1st Difference Intercept and Trend	Result
Retail	-1.30 (2)	-3.01 (2)	-17.73* (1)	-17.74* (1)	I(1)
Wholesale	-1.45 (4)	-3.40*** (4)	-16.43* (3)	-16.42* (3)	I(0)
Farm	-1.99 (4)	-3.24*** (4)	-16.24* (3)	-16.25* (3)	I(0)
KPSS t-statistic	**Levels/Intercept**	**Levels Intercept and Trend**	**1st Difference Intercept**	**1st Difference Intercept and Trend**	**Result**
Retail	3.11*	0.25*	0.14	0.08	I(1)
Wholesale	2.87*	0.26*	0.104	0.103	I(1)
Farm	2.74*	0.214	0.12	0.07	I(0)

Note: Critical values are -3.44, -2.86, and -2.57 for levels and intercept and -3.97, -3.41, and -3.13 for levels, intercept, and trend respectively at 1%, 5% and 10% for ADF. Critical values are 0.74, 0.46, and 0.35 for levels and intercept and 0.22, 0.15, and 0.12 for levels, intercept, and trend respectively at 1%, 5% and 10% for KPSS. *, **, *** indicate rejection of the null hypothesis at 1%, 5% and 10% levels, respectively.

Table 4.3 Zivot-Andrews Unit Root Test Allowing for One Structural Break

t-statistic (BIC lag)-Break Point	Levels/ Intercept	Levels/ Intercept and Trend	1st Difference/ Intercept	1st Difference / Intercept and Trend	Result Intercept	Result Intercept and Trend
Retail	-3.83 (2) 1993m6	-3.96 (2) 1982m7	-17.98* (1) 1979m6	-18.07* (1) 1979m6	I(1)	I(1)
Wholesale	-4.76 (4) 1993m6	-4.70 (4) 1993m6	-16.55* (3) 1979m6	-16.57* (3) 1979m6	I(1)	I(1)
Farm	-4.27 (4) 1993m6	-4.18 (4) 1993m6	-16.37* (3) 1999m1	-16.42* (3) 2015m6	I(1)	I(1)

Note: * indicates rejection of the null hypothesis at 5%. Critical values are -4.80 and -5.08 for intercept and intercept and trend, respectively.

Table 4.4 Granger Causality Test Results

Null hypothesis	F-Statistic
Farm price does not Granger cause wholesale price	19.84*
Wholesale price does not Granger cause farm price	1.43
Wholesale price does not Granger cause retail price	82.08*
Retail price does not Granger cause wholesale price	17.20*
Farm price does not Granger cause retail price	76.55*
Retail price does not Granger cause farm price	23.19*

Note: *1% significance level.

Table 4.5 L+1 vs. L Sequentially Determined Breaks

Break Test	F-statistics	Critical Value**	Number of Breaks	Dates
0 vs. 1*	36.92	13.98	1	1980M11
1 vs. 2*	29.56	15.72	2	1993M07
2 vs. 3*	21.86	16.83	3	2001M05
3 vs. 4*	19.52	17.61	4	2013M09
4 vs. 5	0.00	18.14	4	

Note: * indicates rejection of the null hypothesis at 5%. ** Bai-Perron (2003) critical values

Table 4.6 Johansen Cointegration Test Results

Null Hypothesis[a]	Trace statistics	5% Critical value	Eigenvalue
$r = 0$*	146.12	21.13	0.00
$r \leq 1$*	39.14	14.26	0.00
$r \leq 2$	0.48	3.84	0.49

Note: [a]: r is the cointegrating rank, *rejection of the null hypothesis at the 5% level.

Table 4.7 The Empirical Estimates of Speeds of Adjustment and Diagnostics

Variable	ΔP_{retail}	$\Delta P_{wholesale}$	ΔP_{farm}
Speeds of adjustment	-0.021**	0.206*	0.074*
Model diagnostics			
R^2	0.38	0.23	0.17
AIC	-5.37	-3.25	-3.47
SIC	-5.23	-3.10	-3.33

Note: *1% and **5% significance level.

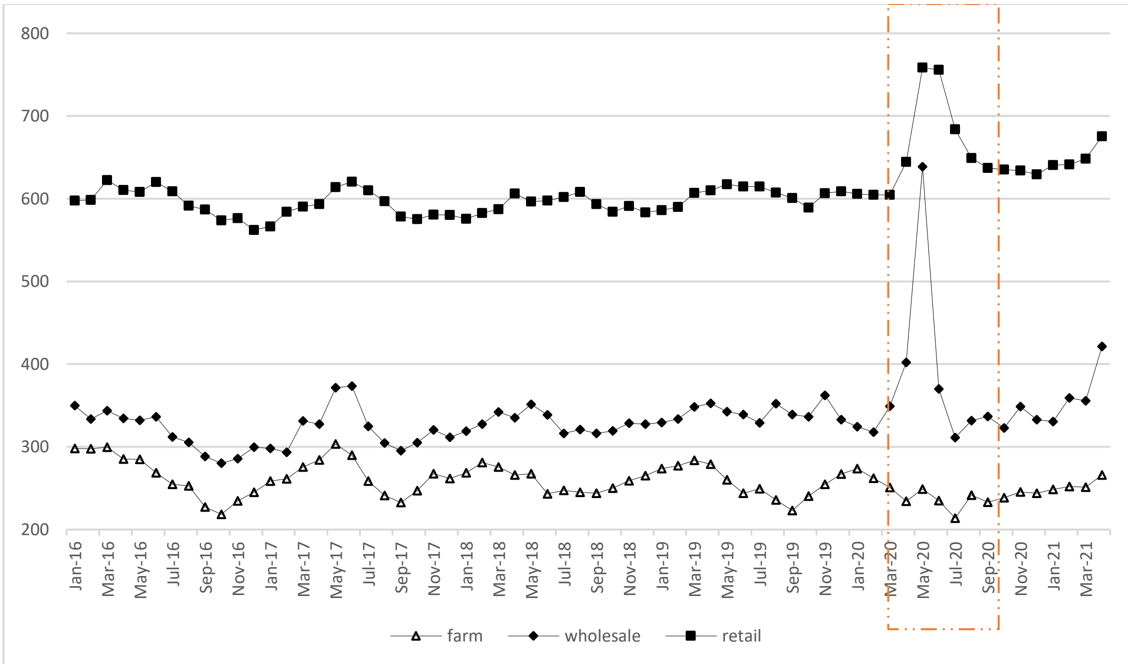

Figure 4.1 U.S. Beef Prices (Cents per Pound)

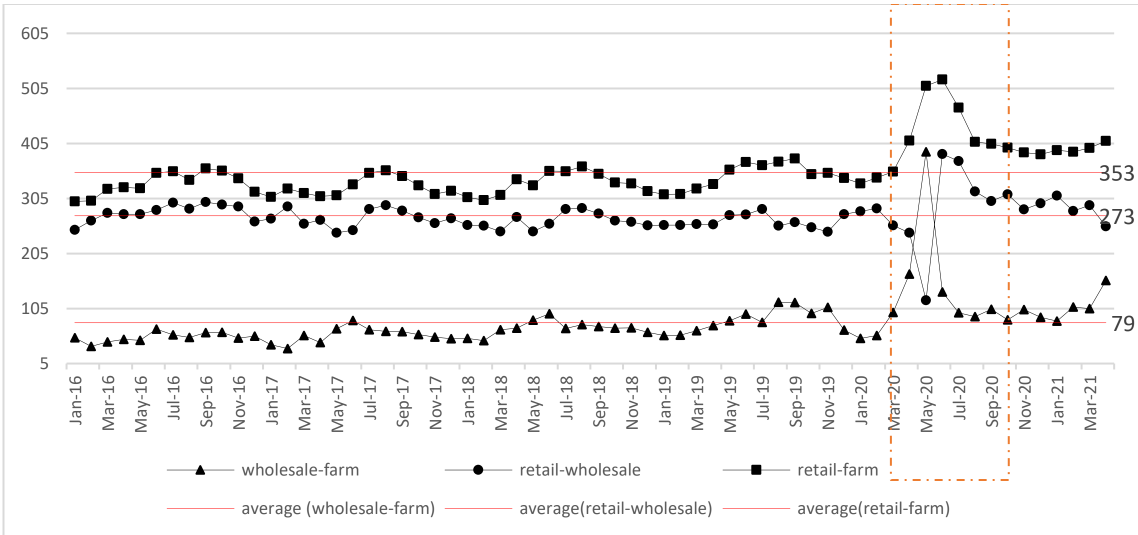

Figure 4.2 U.S. Beef Price Spreads (Cents per Pound - Averages of 2016-2021M04)

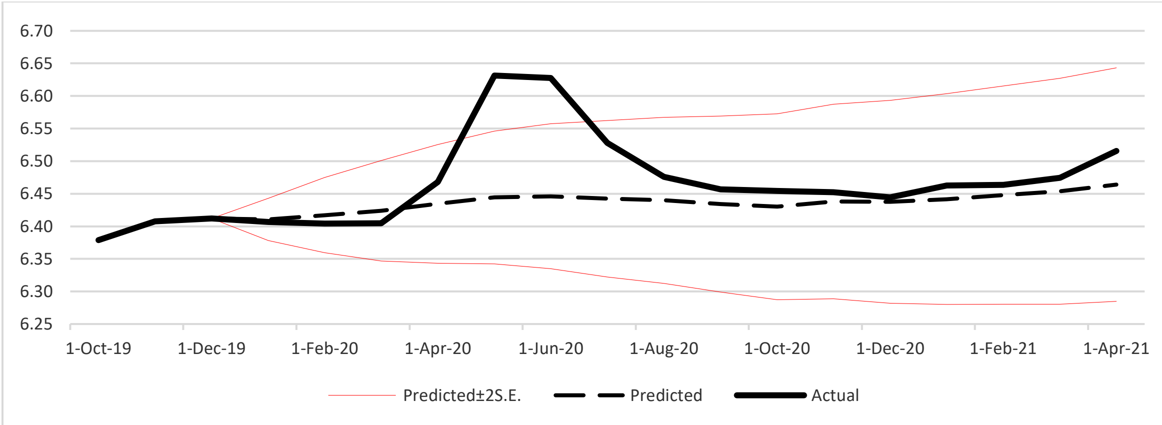

Figure 4.3a Retail Prices

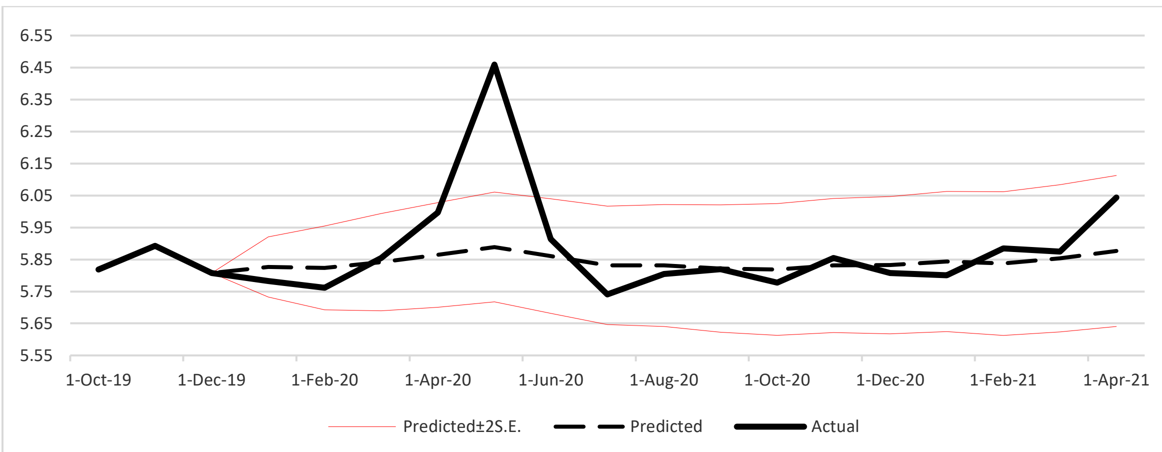

Figure 4.3b Wholesale Prices

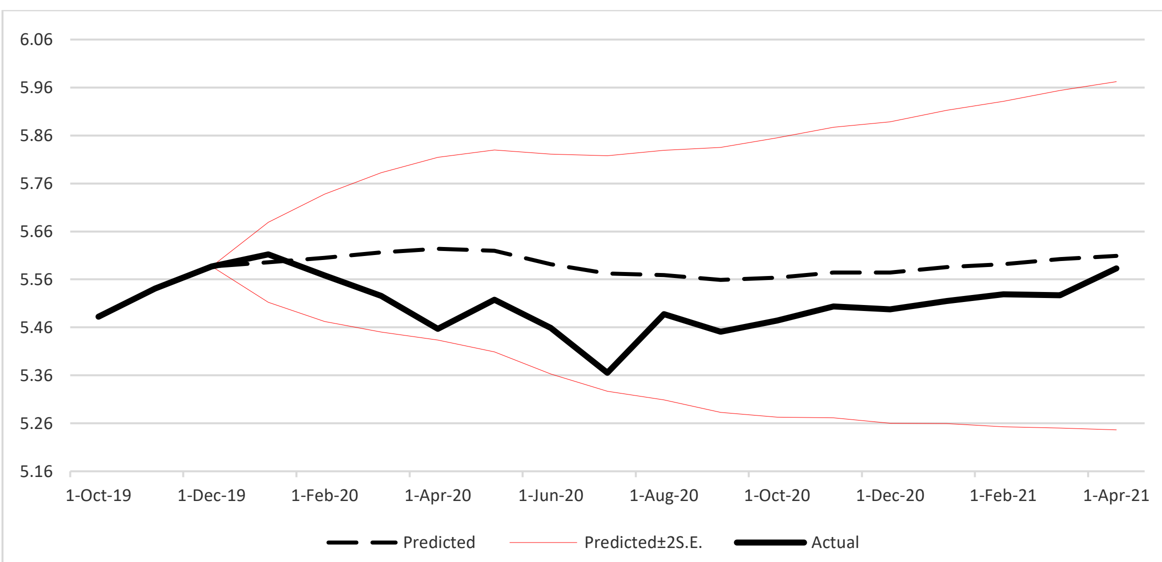

Figure 4.3c Farm Prices

CHAPTER 5. CONCLUSION

This dissertation employs various quantitative methods to address potential challenges in the U.S. beef supply chain and provides empirical results with useful implications and tools for researchers, extension specialists, and stakeholders.

Chapter 2 evaluates optimal culling decision strategies in cow-calf operations in the U.S with a novel dynamic linear programming model. It estimates a base model and runs several experiments with a set of assumptions related to production and replacement heifer costs, cow fertility, calf weights, prices, and pregnancy check use. The empirical analysis is performed with the data obtained for a spring calving herd in Kentucky. The results indicate that producers should cull all cows that are older than age 10 considering their productivity, production costs, and product prices in the base model. The model suggests culling open cows earlier (at age 7) given their productivity status and probabilities. The sensitivity analysis presents a range of optimal decisions that vary with respect to market conditions, cost structure, cow fertility, calf weights, and pregnancy check use. While the change in cow-calf prices considerably affects net return values among the conducted experiments, the cost sensitivity analysis which involves altering bred heifer replacement value, has the most substantial impact on both net return and herd age decomposition. The pregnancy checking experiment results led to an increase in the net return above selected costs and suggest that producer should only retain open cows that are 4-year-old and younger and tested as pregnant.

Chapter 3 proposes a model selection framework which compares traditional time series techniques and machine learning algorithms to provide one-week-ahead steer, heifer, and cull cow cash price forecasts in the southern U.S. The study uses weekly

Kentucky cattle auction prices with lagged variables and dummy variables for weekly seasonal structure. The empirical results of this essay reveal that ARIMA models without seasonality provide the best forecasts for steer prices and the LASSO regression is the best technique to forecast heifer and cull cow prices. The model selection results point to the superiority of machine learning techniques over standard ARIMA models when forecasting U.S. livestock prices in larger samples and suggest that machine learning techniques should be considered as an alternative forecasting tool for extension specialists and producers.

Chapter 4 examines the price dynamics along the U.S. beef supply chain and estimates the impact of COVID-19 on the dynamics of vertical price transmission in the U.S. beef industry using monthly farm, wholesale, and retail prices for the period from 1970 to 2021. The empirical results of this essay provide several implications that contribute to the current literature. The findings indicate that the impact of COVID-19 has been uneven across the beef marketing channel with farmers bearing the burden of the shock. The results also emphasize that in the case of the COVID-19 shock, wholesale prices adjusted more quickly than both farm (threefold) and retail prices (tenfold). Historical decomposition graphs reveal that the COVID-19 pandemic led to higher prices for retailers and wholesalers, while farmers received lower prices than their predicted values. Therefore, farmers in the U.S. beef supply chain were adversely affected by the COVID-19 pandemic. The results point out that the U.S. beef markets were resilient enough to absorb the shocks and return to their pre-shock patterns in 4 to 6 months.

APPENDIX: DATA INTERPOLATION PROCESS (CHAPTER 3)

Missing Data

Missing data statistics are given in Table A.1. The Kalman filter function in R package imputeTS is used to generate missing observations. The selected Kalman filter function fits a state space model on time series and estimates value of missing observation based on available observations before and after the missing value, and it produces good results for time series with strong seasonality and trend (Moritz & Bartz-Beielstein, 2017).

KY Price Predictions for the Period 1999-2010

VAR causality tests and Structural VAR (SVAR) models are employed to obtain models to predict the statistical relation between KY and KS series for the period 2010-2020. Using these models, KY price series are generated for the period 1999-2009.

VAR Granger Causality Test

The system in A.1 is used to construct a VAR model for steer and heifer prices. The optimal lag length for both variables is determined to be two based on BIC in lag selection.

$$KY_t = \alpha + \beta_{11}KY_{t-1} + \beta_{12}KY_{t-2} + \beta_{13}KS_{t-1} + \beta_{14}KS_{t-2} + u_{kyt}$$

$$KS_t = \delta + \beta_{21}KY_{t-1} + \beta_{22}KY_{t-2} + \beta_{23}KS_{t-1} + \beta_{24}KS_{t-2} + u_{kst} \tag{A.1}$$

Here KY_t and KS_t are the price series, α and δ are intercepts, β coefficients are short-run dynamic coefficients of the model's adjustment to the long-run, and u_{kyt} and u_{kst} are residuals in the equation.

The results of all three tests in Table A.2 and A.3 confirm that KS prices cause KY prices.

SVAR Model

After confirming the causal relation between series, a SVAR model is estimated for steer and heifer prices to find the coefficients for the contemporary and lagged relations between series. The system in A.2 is employed to construct the SVAR model and the optimal lag length is determined to be two for both variables based on BIC.

$$KY_t + a_{12}KS_t = \alpha + \beta_{11}KY_{t-1} + \beta_{12}KY_{t-2} + \beta_{13}KS_{t-1} + \beta_{14}KS_{t-2} + u_{kyt}$$

$$a_{21}KY_t + KS_t = \delta + \beta_{21}KY_{t-1} + \beta_{22}KY_{t-2} + \beta_{23}KS_{t-1} + \beta_{24}KS_{t-2} + u_{kst} \tag{A.2}$$

In the system, KY_t and KS_t are the price series, α and δ are intercepts, β coefficients determine short-run relations, and u_{kyt} and u_{kst} are residual in the equation. a_{12} and a_{21} are the main coefficients of interest as they represent contemporary relation between price series.

It is assumed that KY prices do not impact KS prices contemporary and the restriction $a_{21} = 0$ is imposed in the model as identification strategy to make the system recursive (Wold, 1951). Estimated coefficients for steer and heifer prices are presented in Table A.4. These are the best models to predict KY prices with KS prices for the period 2010-2020. These coefficients are used to generate a KY sample for the period 6/11/1999-01/01/2010.

The average difference (△) between series for the period 01/08/2010-12/18/2020 is used to compute the first and second observations.

$$\triangle = E[KY_t^{2010-2020} - KS_t^{2010-2020}]$$

$$KY_{t=1} = KS_{t=1} + \triangle$$

$$KY_{t=2} = KS_{t=2} + \triangle$$

The following formulation is used to compute the remaining observations starting from $t = 3$.

Steer prices:

$$KY_t = -1.925 + 0.285 * KS_t + 0.666 * KY_{t-1} + 0.116 * KY_{t-2} + 0.0448 * KS_{t-1} - 0.123 * KS_{t-2} \qquad \text{(A.3)}$$

Heifer prices:

$$KY_t = -0.531 + 0.285 * KS_t + 0.762 * KY_{t-1} + 0.130 * KY_{t-2} - 0.002 * KS_{t-1} - 0.183 * KS_{t-2} \qquad \text{(A.4)}$$

Predicted values for 6/11/1999-01/01/2010 and actual values for 01/08/2010-12/18/2020 are combined and used to forecast prices.

Table A.1 Missing Data Statistics

Variables	Description	Sample Period	Sample Size	Missing Observations	
				Number	Percentage (%)
KYS	KY Steer Prices	01/08/2010-12/19/2020	572	41	7.17
KYH	KY Heifer Prices	01/08/2010-12/19/2020	572	45	7.87
KSS	KS Steer Prices	6/11/1999-12/18/2020	1,124	92	8.19
KSH	KS Heifer Prices	6/11/1999-12/18/2020	1,124	100	8.90
KYC	KY Cull Cow Prices	2/5/1993-12/19/2020	1,455	113	7.77

Table A.2 VAR Models and t-statistic Test Results for KY Steer and Heifer Prices

Coefficients	KY Steer	KY Heifer
α	-0.714	0.334
	(0.975)	(0.819)
β_{11}	0.769*	0.859*
	(0.044)	(0.0438)
β_{12}	0.0706	0.0710
	(0.0438)	(0.044)
β_{13}	0.237*	0.180*
	(0.0373)	(0.035)
β_{14}	-0.0885***	-0.119*
	(0.0375)	(0.0345)
Observations	570	570

Note: *, **, *** indicate statistical significance at the 0.01, 0.05 and 0.10 level, respectively. Standard errors in parentheses.

Table A.3 VAR Granger Causality Test

	$H_0: There\ is\ no\ causal\ relationship\ between\ the\ series$ i.e., KS do not Granger-cause KY	
	Chi^2	Probability> Chi^2
KYS	57.17	0.000
KYH	63.59	0.000

Note: KYS: Kentucky Steer Prices, KYH: Kentucky Heifer Prices.

Table A.4. SVAR Models for KY Steer and Heifer Prices

Coefficients	KY Steer	KY Heifer
α	-1.925***	-0.531
	(0.931)	(0.773)
a_{12}	0.284*	0.285*
	(0.0339)	(0.0317)
β_{11}	0.666*	0.762*
	(0.0433)	(0.0424)
β_{12}	0.116**	0.130**
	(0.0416)	(0.0417)
β_{13}	0.0448	-0.002
	(0.0419)	(0.038)
β_{14}	-0.123*	-0.183*
	(0.0357)	(0.0329)

Note: *, **, *** indicate statistical significance at the 0.01, 0.05 and 0.10 level, respectively. Standard errors in parentheses.

REFERENCES

Adhikari, R., & Agrawal, R. K. (2013). *An introductory study on time series modeling and forecasting*. Lambert Academic Publishing.

Amadou, Z., Curry Raper, K., Biermacher, J. T., Cook, B., & Ward, C. E. (2014). Net returns from feeding cull beef cows: The influence of initial body condition score. *Journal of Agricultural and Applied Economics*, *46*(1), 139–155.

Arnold, M., Burris, R., & Townsend, L. (2021). Health and management techniques. In L. Anderson (Ed.), *The Kentucky Beef Book* (1st ed., pp. 93–118). University of Kentucky College of Agriculture, Food and Environment, Cooperative Extension Service.

Ates, A. M., Lusk, J. L., & Brorsen, B. W. (2019). Forecasting meat prices using consumer expectations from the Food Demand Survey (FooDS). *Journal of Food Distribution Research*, *50*(1).

Azzam, S. M., & Azzam, A. M. (1991). A Markovian decision model for beef cattle replacement that considers spring and fall calving. *Journal of Animal Science*, *69*(6), 2329–2341.

Bai, J. (1997). Estimation of a change point in multiple regression models. *The Review of Economics and Statistics*, *79*(4), 551–563.

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, *18*(1), 1–22.

Bairoliya, N., & İmrohoroğlu, A. (2020). Macroeconomic consequences of stay-at-home policies during the COVID-19 pandemic. *Covid Economics: Vetted and Real-Time Papers*, *13*.

Balagtas, J. v., & Cooper, J. (2021). The impact of COVID-19 on United States meat and livestock markets. *AAEA Choices Magazine*, *36*(3).

Baldwin, R., & di Mauro, B. W. (2020). *Economics in the time of COVID-19*. CEPR Press.

Barman, A., Das, R., & De, P. K. (2021). Impact of COVID-19 in food supply chain: Disruptions and recovery strategy. *Current Research in Behavioral Sciences*, *2*, 100017.

Beckman, J., & Countryman, A. M. (2021). The importance of agriculture in the economy: Impacts from COVID-19. *American Journal of Agricultural Economics*, *103*(5), 1595–1611.

Bentley, E., & Shumway, C. R. (1981). Adaptive planning over the cattle price cycle. *Southern Journal of Agricultural and Applied Economics*, *13*(1), 139–148.

Bentley, E., Waters, J. R., & Shumway, C. R. (1976). Determining optimal replacement age of beef cows in the presence of stochastic elements. *Journal of Agricultural and Applied Economics*, *8*(2), 13–18.

Bergmeir, C., Hyndman, R. J., & Koo, B. (2015). A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Monash Econometrics and Business Statistics Working Paper*, *10:15*.

BIF, Beef Improvement Federation. (2018). *Guidelines for uniform beef improvement programs* (Ninth Edition-Rev.).

Blevins, P. (2009). Marketing cull cows in Virginia. Virginia Cooperative Extension Service, E-400- 761, Communications and Marketing, College of Agriculture and Life Sciences, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 2009.

Boetel, B. L., & Liu, D. J. (2010). Estimating structural changes in the vertical price relationships in U.S. beef and pork markets. *Journal of Agricultural and Resource Economics*, *35*(2), 1–17.

Boyer, C. N., Griffith, A. P., & DeLong, K. L. (2020). Reproductive failure and long-term profitability of spring- and fall-calving beef cows. *Journal of Agricultural and Resource Economics*, *45*(1), 78–91.

Bunge, J. (2020, April 6). *Coronavirus hits meat plants as some workers get sick, others stay home*. The Wall Street Journal. https://www.wsj.com/articles/coronavirus-hits-meat-plants-as-some-workers-get-sick-others-stay-home-11586196511

Burdine, K. H., Maynard, L. J., Halich, G. S., & Lehmkuhler, J. (2014). Changing market dynamics and value-added premiums in southeastern feeder cattle markets. *The Professional Animal Scientist*, *30*, 354–361.

Burt, O. R. (1965). Optimal replacement under risk. *Journal of Farm Economics*, *47*(2), 324–346.

Cabrera, Victor. E. (2010). A large Markovian linear program to optimize replacement policies and dairy herd net income for diets and nitrogen excretion. *Journal of Dairy Science*, *93*(1), 394–406.

Cerqueira, V., Torgo, L., & Soares, C. (2019). Machine learning vs statistical methods for time series forecasting: Size matters. *ArXiv*, *1909.13316*.

Chopra, A., & Bessler, D. A. (2005). Impact of BSE and FMD on beef industry in U.K. *NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management*.

Cochrane, C. (2018, May 18). *Time series nested cross-validation.* Towards Data Science. https://towardsdatascience.com/time-series-nested-cross-validation-76adba623eb9

Corah, L., & Lusby, K. (2000). Factor Influencing Conception Rate. In *Beef Cattle Handbook*.

Cowley, C. (2020, July 31). *COVID-19 disruptions in the U.S. meat supply chain*. Main Street Views: Policy Insights from the Kansas City Fed.

https://www.kansascityfed.org/agriculture/ag-outlooks/COVID-19-US-Meat-Supply-Chain/

Darbandi, E., & Saghaian, S. H. (2016). Vertical price transmission in the U.S. beef markets with a focus on the Great Recession. *Journal of Agribusiness*, *34*(2), 91–110.

Darekar, A., & Reddy, A. A. (2017). Cotton price forecasting in major producing states. *Economic Affairs*, *62*(3), 1–6.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*(366a), 427–431.

Divina, F., Torres, M. G., Vela, F. A. G., & Noguera, J. L. V. (2019). A comparative study of time series forecasting methods for short term electric energy consumption prediction in smart buildings. *Energies*, *12*(10).

Dudek, G. (2015). Short-term load forecasting using random forests. *Advances in Intelligent Systems and Computing*, *323*, 821–828.

Emmanouilides, C. J., & Fousekis, P. (2015). Vertical price dependence structures: Copula-based evidence from the beef supply chain in the USA. *European Review of Agricultural Economics*, *42*(1), 77–97.

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, *55*(2), 251–276.

English, L., Popp, J., Alward, G., & Thoma, G. (2020). *Economic contributions of the US beef industry*. Presented at the National Cattlemen's Beef Association Meeting, 16–17 November 2020.

Erol, E., & Saghaian S. H. (2022). The COVID-19 pandemic and dynamics of price adjustment in the U.S. beef sector. *Sustainability*, *14(8), 4391*.

Erol, E. (2022, April 19). *Overview of the U.S. beef industry*. https://erdalerol.github.io/US-Beef-Industry/

Flores, H., Meneses, C., Villalobos, J. R., & Sanchez, O. (2017). Improvement of feedlot operations through statistical learning and business analytics tools. *Computers and Electronics in Agriculture*, *143*, 273–285.

Fousekis, P., Katrakilidis, C., & Trachanas, E. (2016). Vertical price transmission in the US beef sector: Evidence from the nonlinear ARDL model. *Economic Modelling*, *52*, 499–506.

Frasier, W. M., & Pfeiffer, G. H. (1994). Optimal replacement and management policies for beef cows. *American Journal of Agricultural Economics*, *76*(4), 847–858.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, *29*(5), 1189–1232.

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis*, *38*(4), 367–378.

Godahewa, R., Bergmeir, C., Webb, G. I., & Montero-Manso, P. (2020). A strong baseline for weekly time series forecasting. *ArXiv*.

Goodwin, B. K. (1992). Forecasting cattle prices in the presence of structural change. *Southern Journal of Agricultural Economics*, *24*(2), 11–22.

Goodwin, B. K., & Holt, M. T. (1999). Price transmission and asymmetric adjustment in the U.S. beef sector. *American Journal of Agricultural Economics*, *81*(3), 630–637.

Guney, S. (2015). An evaluation of price forecasts of the cattle market under structural changes. Proceedings of the 2015 AAEA&WAEA Joint Annual Meeting, July 26-28, San Francisco, California.

Halich, G., Burdine, K., & Shepherd, J. (2022). Cow-calf profitability estimates for 2021 and 2022 (Spring calving herd). *Economic & Policy Update*.

Harris, J. (Jay). (2017). *A machine learning approach to forecasting food prices* [Master]. Dalhousie University.

Harris, K. S., & Leuthold, R. M. (1985). A comparison of alternative forecasting techniques for livestock prices: A case study. *North Central Journal of Agricultural Economics*, *7*(1), 40–50.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning data mining, inference, and prediction* (2nd ed.). Springer Science & Business Media.

Helmers, G. A., & Held, L. J. (1977). Comparison of livestock price forecasting using simple techniques, forward pricing and outlook information. *Western Journal of Agricultural Economics*, *1*(1), 157–160.

Hersom, M., Thrift, T., & Yelich, J. (2018). *Culling and replacement rate in the beef cow herd*. https://edis.ifas.ufl.edu/publication/AN323

Ibendahl, G. A., Anderson, J. D., & Anderson, L. H. (2004). Deciding when to replace an open beef cow. *Agricultural Finance Review*, *64*(1), 61–74.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2017). *An introduction to statistical learning*. Springer Science & Business Media.

Jha, G. K., & Sinha, K. (2013). Agricultural price forecasting using neural network model: An innovative information delivery system. *Agricultural Economics Research Review*, *26*(2), 229–239.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, *59*(6), 1551-1580.

Kastens, T. L., Jones, R., & Schroeder, T. C. (1998). Futures-based price forecasts for agricultural producers and businesses. *Journal of Agricultural and Resource Economics*, *23*(1), 294–307.

Kohzadi, N., Boyd, M. S., Kermanshahi, B., & Kaastra, I. (1996). A comparison of artificial neural network and time series models for forecasting commodity prices. *Neurocomputing*, *10*(2), 169–181.

Kurumatani, K. (2020). Time series forecasting of agricultural product prices based on recurrent neural networks and its evaluation method. *SN Applied Sciences*, *2*(8).

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, *54*(1–3), 159–178.

Laborde, D., Martin, W., & Vos, R. (2021). Impacts of COVID-19 on global poverty, food security, and diets: Insights from global model scenario analysis. *Agricultural Economics*, *52*(3), 375–390.

Lehenbauer, T. W., & Oltjen, J. W. (1998). Dairy cow culling strategies: making economical culling decisions. *Journal of Dairy Science*, *81*(1), 264–271.

Livanis, G. T., & Moss, C. B. (2005). Price transmission and food scares in the U.S. beef sector. *2005 American Agricultural Economics Association Annual Meeting*.

Lloyd, T. (2017). Forty years of price transmission research in the food industry: Insights, challenges and prospects. *Journal of Agricultural Economics*, *68*(1), 3–21.

Mackay, W. S., Whittier, J. C., Field, T. G., Umberger, W. J., Teichert, R. B., & Feuz, D. M. (2004). To replace or not to replace: Determining optimal replacement rates in beef cattle operations. *The Professional Animal Scientist*, *20*(1), 87–93.

Maki, W. R. (1963). Forecasting livestock supplies and prices with an econometric model. *Journal of Farm Economics*, *45*(3), 612–624.

Martinez, C. C., Boyer, C. N., & Burdine, K. H. (2021). Price determinants for feeder cattle in Tennessee. *Journal of Agricultural and Applied Economics*, *53*(4), 552–562.

Masini, R. P., Medeiros, M. C., & Mendes, E. F. (2020). Machine learning advances for time series forecasting. *ArXiv*.

McCarl, B. A., & Spreen, T. H. (1997). *Applied mathematical programming using algebraic systems*.

Mead, D., Ransom, K., Reed, S. B., & Sager, S. (2020). The impact of the COVID-19 pandemic on food price indexes and data collection. *Monthly Labor Review*, 1–13.

Melton, B. E. (1980). Economics of beef cow culling and replacement decisions under genetic progress. *Western Journal of Agricultural Economics*, *5*(2), 137–147.

Meyer, J., & von Cramon-Taubadel, S. (2004). Asymmetric price transmission: A survey. *Journal of Agricultural Economics*, *55*(3), 581–611.

Moorey, S. E., & Biase, F. H. (2020). Beef heifer fertility: Importance of management practices and technological advancements. In *Journal of Animal Science and Biotechnology* (Vol. 11, Issue 1). BioMed Central Ltd.

Moritz, S., & Bartz-Beielstein, T. (2017). imputeTS: Time series missing value imputation in R. *The R Journal*, *9*(1), 207–218.

Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, *31*(2), 87–106.

Naing, Y. W. N., & Htike, Z. Z. (2015). Forecasting of monthly temperature variations using random forests. *ARPN Journal of Engineering and Applied Sciences*, *10*(21).

National Cattlemen's Beef Association. (2016). *National market cow and bull beef quality audit: Beef edition*.

National Institute of Allergy and Infectious Disease. (2021). *Coronaviruses diseases*. https://www.niaid.nih.gov/diseases-conditions/coronaviruses

Özertan, G., Saghaian, S., & Tekgüç, H. (2014). Market power in the poultry sector in Turkey. *Boğaziçi Journal Review of Social, Economic and Administrative Studies*, *28*(2), 19–32.

Pala, A. (2013). Structural breaks, cointegration, and causality by VECM analysis of crude oil and food price. *International Journal of Energy Economics and Policy*, *3*(3), 238–246.

Pan, B., & Yang, Y. (2017). Forecasting destination weekly hotel occupancy with big data. *Journal of Travel Research*, *56*(7), 957–970.

Parish, J. A., Williams, B. R., Coatney, K. T., Best, T. F., & Stewart, C. O. (2018). A hedonic analysis of sale lot traits affecting calf prices in Mississippi auction markets. *Professional Animal Scientist*, *34*(3), 240–249.

Payne, N. D., Karali, B., & Dorfman, J. H. (2019). Can cattle basis forecasts be improved? A Bayesian model averaging approach. *Journal of Agricultural and Applied Economics*, *51*(2), 249–266.

Peel, D., & Meyer, S. (2002, March). Cattle price seasonality. *Managing for Today's Cattle Market and Beyond*.

Peel, D. S., & Doye, D. (2017). *Cull cow grazing and marketing opportunities*. https://extension.okstate.edu/fact-sheets/cull-cow-grazing-and-marketing-opportunities.html

Perrin, R. K. (1972). Asset replacement principles. *American Journal of Agricultural Economics*, *54*(1), 60–67.

Poudel, B. P., Poudel, M. R., Gautam, A., Phuyal, S., Tiwari, C. K., Bashyal, N., & Bashyal, S. (2020). COVID-19 and its global impact on food and agriculture. *Journal of Biology and Today's World*, *9*(5), 221.

Pouliot, S., & Shulz, L. L. (2016). Measuring price spreads in red meat. *Agricultural Policy Review*, *Winter*.

Pozo, V. F., Bachmeier, L. J., & Schroeder, T. C. (2021). Are there price asymmetries in the U.S. beef market? *Journal of Commodity Markets*, *21*, 100–127.

Ramsey, A. F., Goodwin, B. K., Hahn, W. F., & Holt, M. T. (2021). Impacts of COVID-19 and price transmission in U.S. meat markets. *Agricultural Economics*, *52*(3), 441–458.

Regression Analysis of Time Series. (2010). *Regression analysis of time series user's guide*. Estima.

Ridgeway, G. (2020). Generalized boosted models: A guide to the gbm package.

Rojas, C., Andino, A., & Purcell, W. (2008). Retailers' response to wholesale price changes: New evidence from scanner-based quantity-weighted beef prices. *Agribusiness*, *24*(1), 1–15.

Rusiman, M. S., Hau, O. C., Abdullah, A. W., Sufahani, S. F., & Azmi, N. A. (2017). An analysis of time series for the prediction of Barramundi (Ikan Siakap) price in Malaysia. *Far East Journal of Mathematical Sciences*, *102*(9), 2081–2093.

Sabu, K. M., & Kumar, T. K. M. (2020). Predictive analytics in agriculture: Forecasting prices of Arecanuts in Kerala. *Procedia Computer Science*, *171*, 699–708.

Saghaian, S. H. (2007). Beef safety shocks and dynamics of vertical price adjustment: The case of BSE discovery in the U.S. beef sector. *Agribusiness*, *23*(3), 333–348.

Schnaubelt, M. (2019). A comparison of machine learning model validation schemes for non-stationary time series data. Working Paper 11/2019. FAU Discussion Papers in Economics.

Siche, R. (2020). What is the impact of COVID-19 disease on agriculture? *Scientia Agropecuaria*, *11*(1), 3–6.

Storm, H., Baylis, K., & Heckelei, T. (2020). Machine learning in agricultural and applied economics. *European Review of Agricultural Economics*, *47*(3), 849–892.

Strohbehn, D. R. (1994). *A three-year summary of the ISU-IRM-SPA beef cow business record*. Ames.

Sun, C., Shrivastava, A., Singh, S., & Gupta, A. (2017). Revisiting unreasonable effectiveness of data in deep learning era. *ICCV*.

Surathkal, P., Chung, C., & Han, S. (2014). Asymmetric adjustments in vertical price transmission in the US beef sector: Testing for differences among product cuts and quality grade. *2014 Agricultural and Applied Economics Association Annual Meeting, July 27-29, 2014, Minneapolis, Minnesota*.

Swinnen, J., & Vos, R. (2021). COVID-19 and impacts on global food systems and household welfare: Introduction to a special issue. *Agricultural Economics*, *52*(3), 365–374.

Tashman, L. J. (2000). Out-of-sample tests of forecasting accuracy: an analysis and review. In *International Journal of Forecasting* (Vol. 16).

Ticlavilca, A. M., Feuz, D. M., & Mckee, M. (2010). Forecasting agricultural commodity prices using multivariate Bayesian machine learning regression. *Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO.*

Trapp, J. N. (1986). Investment and disinvestment principles with nonconstant prices and varying firm size applied to beef-breeding herds. *American Journal of Agricultural Economics*, *68*(3), 691–703.

Tronstad, R., & Gum, R. (1994). Cow culling decisions adapted for management with CART. *American Journal of Agricultural Economics*, *76*(2), 237–249.

Tronstad, R., Gum, R., Ray, D., & Rice, R. (1993). Range cow culling: Herd performance. In R. Tronstad, J. Sprinkle, & G. Ruyle (Eds.), *Arizona Ranchers' Management Guide* (pp. 27–31).

Tyralis, H., & Papacharalampous, G. (2017). Variable selection in time series forecasting using random forests. *Algorithms*, *10*(114).

USDA. (2020). *National Agricultural Statistics Service*. https://www.nass.usda.gov/index.php

USDA Economic Research Service. (2022). *Sector at a glance*. https://www.ers.usda.gov/topics/animal-products/cattle-beef/sector-at-a-glance/

USDA Foreign Agricultural Service. (2022). *Livestock and poultry: World markets and trade*. https://public.govdelivery.com/accounts/USDAFAS/subscriber/new

USDA. (2020). Beef 2017, Beef cow-calf management practices in the United States, 2017, report 1. USDA–APHIS–VS–CEAH–NAHMS. Fort Collins, CO.

Van Arendok, J. A. M., & Dijkhuizen, A. A. (1985). Studies on the replacement policies in dairy cattle. III. Influence of variation in reproduction and production. *Livestock Production Science*, *13*, 333–349.

Van Arendonk, J. A. M. (1986). Studies on the replacement policies in dairy cattle. IV. Influence of seasonal variation in performance and prices. In *Livestock Production Science (Vol. 14).*

Van Hoyweghen, K., Fabry, A., Feyaerts, H., Wade, I., & Maertens, M. (2021). Resilience of global and local value chains to the Covid-19 pandemic: Survey evidence from vegetable value chains in Senegal. *Agricultural Economics*, *52*(3), 423–440.

Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics*, *7*(1), 1–8.

Vavra, P., & Goodwin, B. K. (2005). Analysis of price transmission along the food chain. *OECD Food, Agriculture and Fisheries Papers*, *3*.

Wang, B., Liu, P., Chao, Z., Junmei, W., Chen, W., Cao, N., O'Hare, G. M. P., & Wen, F. (2018). Research on hybrid model of garlic short-term price forecasting based on big data. *Computers, Materials and Continua*, *57*(2), 283–296.

Ward, H., & Powell, J. (2017). *Culling the beef cow herd*. https://beef.unl.edu/cattleproduction/cullingstrategies#:~:text=Once%20the%20calf%20is%20weaned,end%20of%20the%20breeding%20season.

Williams, G. S., Raper, K. C., DeVuyst, E. A., Peel, D., & McKinney, D. (2012). Determinants of price differentials in Oklahoma value-added feeder cattle auctions. *Journal of Agricultural and Resource Economics*, *37*(1), 114–127.

Wold, H. O. A. (1951). Dynamic systems of the recursive type: Economic and statistical aspects. *Sankhyā: The Indian Journal of Statistics (1933-1960)*, *11*(3/4), 205–216.

Zachariasse, L. C., & Bunte, F. H. J. (2003). How are farmers faring in the changing balance of power along the food chain? *OECD-Conference on Changing Dimensions of the Food Economy: Exploring the Policy Issues.*

Zapata, H. O., & Garcia, P. (1990). Price forecasting with time-series methods and nonstationary data: An application to monthly U.S. cattle prices. *Western Journal of Agricultural Economics123*, *15*(1), 123–132.

Zhu, X., Vondrick, C., Fowlkes, C. C., & Ramanan, D. (2016). Do we need more training data? *International Journal of Computer Vision*, *119*(1), 76–92.

Zivot, E., & Andrews, D. W. K. (1992). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business & Economic Statistics*, *10*(3), 251-270.

VITA

Education

- MSc, Economics, Middle East Technical University, Ankara, Turkey, 1/2017
- BSc, Economics, Middle East Technical University, Ankara, Turkey, 6/2011

Professional Positions

- Graduate Research Assistant, University of Kentucky, Lexington, KY, U.S. 8/2019-5/2023
- Primary Instructor, AEC 309 (International Agriculture, World Food Needs, and U.S. Trade in Agricultural Products), University of Kentucky, Lexington, KY, U.S., Fall 2022
- Teaching Assistant, AEC 309 (International Agriculture, World Food Needs, and U.S. Trade in Agricultural Products), University of Kentucky, Lexington, KY, U.S., Fall 2021
- European Union Expert, Ministry of Agriculture and Forestry, Ankara, Turkey, 1/2012-7/2019

Publications

- Erol, E., and S.H. Saghaian (2022). The COVID-19 Pandemic and Dynamics of Price Adjustment in the U.S. Beef Sector. *Sustainability* 2022, 14, 4391. https://doi.org/10.3390/su14084391

Papers in Progress

- Erol, E., N. Pates, and K. H. Burdine (2023). Forecasting Beef-Cattle Prices in the Southern United States: A Model Selection Framework, Manuscript.
- Erol, E., C. R. Dillon, and K. H. Burdine. (2023). Optimal Beef Cow Culling Strategies in the U.S.: A Dynamic Linear Programming Framework, Manuscript.

Grants, Honors, and Awards

- Graduate Student Travel Grant - Agricultural & Applied Economics Association (AAEA), 2022.
- Dean’s High Honor Roll - Middle East Technical University, 2010-2011
- Dean’s Honor Roll - Middle East Technical University, 2009-2010