IDENTIFYING AGRICULTURAL RETAILERS' GAPS IN UNDERSTANDING OF THE VALUE PROPOSITION FOR LARGE COMMERCIAL PRODUCERS

by

Hailey Utech

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science

Department of Agricultural Economics

West Lafayette, Indiana

May 2022

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Allan Gray, Chair

School of Agricultural Economics

Dr. Brady Brewer

School of Agricultural Economics

Dr. Scott Downey

School of Agricultural Economics

Approved by:

Dr. Nicole Olynk Widmar

I dedicate this thesis to my Grandfather who unquestioningly supported my dreams, challenged me to surpass my goals, and always believed in me. I will be forever grateful for his love and to have had him as my Grandpa.

ACKNOWLEDGMENTS

I would like to thank my committee members for all their time, support, and guidance throughout my research. Each of you were an integral part of the evolution of this thesis and my personal growth during my time at Purdue University. Dr. Gray, thank you for helping me shape my research question, for your patience, and constant support. Dr. Brewer, thank you for always being willing to answer my questions and offering direction through different methods. Dr. Downey, thank you for being willing to be part of my thesis committee and offer support wherever I needed. While not on my committee I also owe a thank you to Lourival Carmo Monaco Neto who offered his guidance and support throughout this project.

I would next like to thank my classmates and friends I have made at my time at Purdue University who have encouraged me and pushed me to be a better Agricultural Economist.

Lastly, I would like to thank my family who has always believed in me and encouraged me to accomplish the things I want.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

In order for agricultural retailers to remain successful in a volatile market, it is imperative that they understand the needs and buying behaviors of their producers. These producers can be divided into four buying segments: the Economic buyer, the Agronomic buyer, the Business buyer, and the Performance buyer by identifying similar buying characteristics. The retailer's ability to correctly predict their producers into the correct buying segment would allow them to optimally market to individual producers offering a consistent value proposition across all farms. This research uses cluster analysis to segment the agricultural market, multinomial logistic regression models to extract the variables that determined cluster classifications, and accuracy measures from a multilevel confusion matrix to assess retailers' ability to classify their producers into the correct buying segment. Retailers predicted 70% of their producers into the correct segment. However, the accuracies differed across each segment leaving opportunity for an inconsistent value proposition across all segments.

CHAPTER 1. INTRODUCTION

1.1 Overview of Farm Consolidation

While small farms make up 90% of the farms and operate half the farmland in the United States, 45% of farm production occurs on large-scale farms (Hoppe, 2017). Along with being more productive, larger farms are also more profitable than smaller farms (Hoppe, 2015). A small farm is much more likely than a large farm to have an operating profit margin less than 10%, indicating high financial risk (Hoppe, 2017). Considering the already low profit margin farmers must work with, it is not surprising that farmers would want to consolidate and have larger farms to take advantage of economies of scale in their operation. The number of farms with at least 2,000 acres of cropland doubled from 20,638 acres in 1987 to 42,620 acres in 2017. There was also a significant increase in farms with cropland over 10,000 acres with 294 farms in 1987 and 1,191 farms in 2017 (Macdonald, 2020).

As farms became more consolidated and inflation-adjusted farm production expenses rose by 4.4% in 2021, agricultural retailers are challenged with how best to differentiate themselves from other retailers, move into and keep the primary supplier position (Farming and Farm Income, 2022). Differentiation can be achieved by either contribution of consumer value and/or by providing a lower cost. However, the optimal balance of these two factors along with the determinants for consumer value being different for different customers result in a demanding situation for retailers to understand and provide the highest value to their producers. Retailers, therefore, must develop an effective marketing strategy to differentiate themselves in a way that will increase their notion of value among producers.

1.2 Problem Statement

Retailers today face many challenges unforeseen in the past. The consolidation of farms, technology, market interruptions, and vastly different customer needs make the understanding of producers' mindsets and perceptions more difficult than ever. Those retailers that can accurately grasp the producers' needs and articulate a compelling value proposition can differentiate themselves from competitors. Retailers differentiating themselves and offering the highest value proposition will be the difference between retailers that continue to grow and succeed and those

who will fall prey to a challenging market. As farms consolidate and become more sophisticated, while retailers' ability to serve greater quantities of farms expands, the need to identify groups of likeminded farmers becomes increasingly more important to allow retailers to maximize their resources and provide a consistent value proposition across all farms.

1.3 Objective

This study examines the gap between agricultural retailer's opinions of the value proposition they offer farmers and the value propositions articulated by farmers that they seek from their agricultural retail service providers in terms of both products and services. The objective of this study is to identify how demographics and other seller attributes explain differences between retailers' and producers' perception of value created for large commercial producers.

1.4 Hypotheses

The following hypotheses are expected based on prior research and market knowledge:

- Hypothesis 1: Significant gaps exist in the buying segments derived from producers expressed buying preferences relative to segments derived from Retailers' perceptions of large producer buying behaviors
- Hypothesis 2: Retailers' perceptions of farmer buying behavior results in significant differences in the sizes of producer buying segments relative to the size of buying segments identified using actual producer data

1.5 Methodology

To test these hypotheses the farmer data from the 2021 edition of the Large Commercial Producer Survey and 2021 edition of the Agricultural Retailer Survey was used. First, cluster analysis using a combination of Ward's Method and K-Means was used to find groups with similar psychographic characteristics illustrated in Figure 1-1.

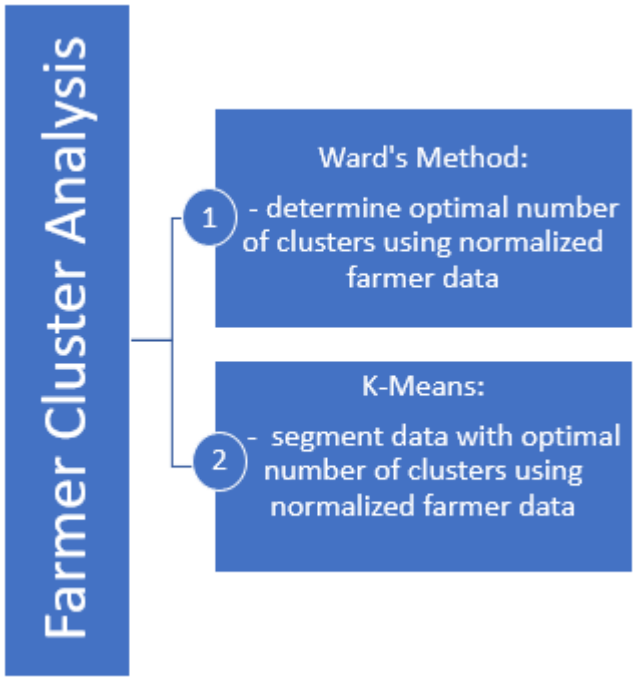

Figure 1-1 Cluster Analysis

Then, a multinomial logistic model was used to predict cluster classification for the farmer and retailer data:

$$\begin{aligned} Cluster_i = {} & Q_{11i} + Q_{12i} + Q_{13i} + Q_{14i} + Q_{15i} + Q_{16i} + Q_{21i} + Q_{22i} + Q_{23i} \\ & + Q_{24i} + Q_{25i} + Q_{26i} + Q_{31i} + Q_{32i} + Q_{33i} + Q_{34i} + Q_{41i} \\ & + Q_{42i} + Q_{43i} + Q_{51i} + Q_{52i} + Q_{53i} + Q_{61i} + Q_{62i} + Q_{63i} \\ & + Q_{64i} + Q_{65i} + Q_{71i} + Q_{72i} + Q_{73i} + Q_{74i} + Q_{81i} + Q_{82i} \\ & + Q_{83i} + Q_{84i} \end{aligned} \quad (1\text{-}1)$$

where Q_{11i} represents the response from the i^{th} observation for segmentation question 1, variable 1.

These multinomial logistic regression models were then used to directly compare farmer cluster classifications predicted by the producer multinomial logistic regression model and farmer cluster classifications predicted by the retailer multinomial logistic regression model. This difference between these classifications will identify the key gaps in retailers' understanding of value proposition for large commercial producers.

1.6 Organization of Thesis

To understand the gap between agricultural retailers' understanding of consumer value proposition for large commercial producers, this thesis will proceed in the following manner. This chapter provided an introduction to consolidation of farms in the United States and the importance of retailer differentiation. Chapter 2 contains a literature review on value proposition theory, segmentation of the agricultural market, and clustering and multinomial logistic model. Chapter 3 follows with the conceptual framework used in this study, along with justifications for methods used. Chapter 4 includes an overview of the data collection process and descriptive statistics for the sample. Chapter 5 contains the results and interpretations from this study. And lastly, chapter 6 includes a summary and study conclusion. There will also be two appendices, the first containing the survey questions used and the second consisting of the tables referenced in the analysis and results chapter.

CHAPTER 2. LITERATURE REVIEW

2.1 Overview

This literature review is a summary and discussion of the literature around value proposition theory, agricultural market segmentation, and the methodology used when segmenting the market. The first section, Value Proposition Theory, defines the creation of value and elaborates on how businesses can increase their value-added strategies. The second section, Segmenting the Agricultural Market, focuses on the previous body of research around segmenting the agricultural market including similarities between studies, how this research has evolved, and results from each study. The third section, Clustering and Multinomial Logistic Model, focuses on the clustering and multinomial logistic model methodology used in the studies in the previous section.

2.2 Value Proposition Theory

Ravald and Grönroos (1996) introduced the concept of total episode value combining both the individual exchange value (episode value), where episode value can be further broken down into episode benefits and episode sacrifices and the value derived from the customer-retailer relationship.

$$Total\ episode\ value = \frac{Episode\ Benefits + Relationship\ Benefits}{Episode\ Sacrifice + Relationship\ Sacrifice} \qquad (2\text{-}1)$$

Formula 2-1 emphasizes the importance of a continuing relationship between the retailer and customer. Value-added strategies add another dimension to customer satisfaction and retention. If the retailer increased the benefits and reduced the sacrifice → stimulation for repurchase → relationship with retailer → credibility, safety, and security within the relationship → trust between retailer and customer → loyalty for purchasing → mutually profitable relationship. Understanding how customers perceive the value-added strategies retailers use is important for optimally increasing the benefits for the customer. Another way to increase the total episode value is decreasing the sacrifice. On the surface, decreasing the sacrifice can look like decreasing the price or making it more convenient by reducing the time it takes to obtain the purchase. However, increasing the benefits while reducing the sacrifice can become a difficult

balancing act for the retailer. A value-added strategy could increase the total episode value for one customer but decrease it for another. Understanding the magnitude these value-added strategies have on customer purchasing decisions and retention can lead to mutual benefits for the customer and retailer.

Ulaga and Eggert (2006) studied how companies can differentiate themselves to achieve key supplier status through value creation for business-to-business relationships. The authors define customer perceived value in a key supplier relationship as "a higher-order construct that represents the trade-off between the benefits and the costs perceived in the supplier's core offerings, in the sourcing process, and at the level of a customer's operations, taking into consideration the available alternative supplier relationships." This framework builds on Ravald and Grönroos' (1996) definition of how value-added strategies can impact total episode value by taking into consideration relationships with other suppliers. The authors gathered data from purchasing managers in manufacturing companies across the Midwest in the United States. The respondents for this data consisted of manufacturers from a variety of areas including aircraft landing systems, speaker equipment, automobiles, and household appliances. The results of this study suggest relationship benefits have a stronger impact on key supplier differentiation than cost considerations. Respondents identified service support and personal interaction as the most important factors of differentiation followed by the supplier's knowledge and ability to reduce manufacturing time. Product quality, delivery performance, acquisition cost, and operation costs had a moderate impact on acquiring and maintaining key supplier status. Price had the weakest impact for key supplier differentiation.

This research offers a conceptual framework for the role of value in consumers' decision-making process. The importance of product knowledge, market knowledge, and innovation will be analyzed with respect to large commercial producers that agricultural retailers serve. Analyzing large commercial producers' perspectives in these areas will help determine where retailers may be missing the producer's value proposition in their services.

Payne, Frow, and Eggert (2017) examine three different conceptual perspectives of consumer value proposition. The supplier-determined consumer value proposition views customer value through a supplier value delivery system comprised of three key stages: choose, provide, and communicate the value proposition. The transitional consumer value proposition emphasizes the understanding of customers' perspectives and experiences during

use of the product or service that brings this perspective closer to a value-in-use perspective. The mutually determined consumer value proposition allows for co-creation of value between the retailer and consumer.

The authors proposed a working definition of the consumer value proposition, "a customer value proposition (CVP) is a strategic tool facilitating communication of an organization's ability to share resources and offer a superior value package to targeted customers." This definition, in comparison to previous definitions, highlights consumer value propositions communication role, accentuates the role of resource sharing between the retailer and the consumer, and highlights the importance of segmentation by providing targeted packages of value to consumers.

Payne, Frow, and Eggert (2017) believe that managing customer perceptions over time is highly relevant to providing value. The theory behind managing customer perceptions can be applied when looking at agricultural producers' responses as to whether they perceive the value the retailer intended.

Eggert, Ulaga, Frow, and Payne (2018) studied how value is created and has evolved from a focus on resource exchange and value in exchange, meaning the supplier creates value through products exchanged with customers, to an emphasis on value in use, meaning value is jointly created by the supplier and their customer. The evolution of these different value perspectives is comprised of three stages. The primary stage conceptualizes value from the customers' perspective as being a trade-off between quality and price. The second stage focuses on expanding understanding of value perceptions in business relationships and value creation in customers' use situations. The final state uses consumer value proposition as a strategic tool for businesses to convey value to their customers. Communicating value aimed at the customer under the value in exchange framework allows value to be viewed as co-created between the business and customer.

Baker (2006) argues that customer value is determined not only by the traditional elements of history, reputation/brand name, quality, trust, and commitment but also experience. The author found that "only 15 to 35 percent of customers consider price to be the chief determinant, depending on what they are buying…greater than 60 percent do not consider price at all, and almost 80 percent cannot correctly recall (within 10 percent) the price they paid for a product in the past seven days, although they do remember the brand" (Baker, 2006). However, 65 percent of customers prefer value to low prices. Following Baker's findings, our study will examine large

commercial producers' preferences for prices relative to value as compared to retailers' perceptions of producers' preferences.

2.3 Segmentation of the Agricultural Market

In this section, previous research around the segmentation of agricultural markets will be discussed. This research provided background and prior producer segmentation which was taken into consideration for this research.

Gloy and Akridge (1999) studied the segmentation of the commercial producer marketplace for agricultural inputs by using cluster analysis. There were twelve questions that served as the basis for segmentation clustering variables. In this research age, education, and other demographic characteristics were left out when deriving the segments with the belief that demographic data tends to not be directly related to buying preferences although it may be indirectly related. The segments found in this research represent the difference desires commercial producers have for their input suppliers. These producers can be split into four segments: Balance, Convenience, Performance, and Price buyers.

The Balance segment represented the largest proportion of producers and led to the prediction that producers in this segment place value on an input supplier that can provide a wide array of services, information, reasonable prices along with products that perform well. Producers in the Convenience segment placed a large importance on convenience and location factors. Producers in the Performance segment placed value on products that work well and are like the Balance segment in how they value price as a factor of importance. Producers in the Price segment placed a very low importance on personal factors and support services which indicates that these producers likely change suppliers often.

Alexander, Wilson, and Foley (2005) studied the agricultural commercial producers' market by segmenting producers into five main market segments of balance, price, performance, convenience, and service. The data used for this research is from the 1998 and 2003 editions of the Commercial Producer Projects conducted by the Center for Food and Agricultural Business at Purdue University. Alexander, Wilson, and Foley (2005) expanded on Gloy and Akridge (1999) by proposing the use of a multinomial logit model to predict the segmentation of agricultural producers through demographic variables after the segmentation was complete.

Alexander et al. (2009) used the 2008 edition of the Large Commercial Producer Survey conducted by the Center for Food and Agricultural Business at Purdue University to study effective and efficient communication depending on operation type and demographics of agricultural producers. Producers were segmented based on their response to the question of how convenience/location, service/information, price, product performance, and support services influenced their purchase of seed, crop protection chemicals, animal health, feed, capital equipment, and financial products. Similar to Gloy and Akridge (1999), four segments: Balance, Convenience, Performance, and Price were found for expendable products such as seed, crop protection chemicals, animal health, and feed.

Harbor, Black, Babin, and Akridge (2008) uses the survey data from the 2003 edition of the Large Commercial Producer Survey conducted by the Center for Food and Agricultural Business at Purdue University to assess the prevalence and factors of brand loyalty for agricultural inputs. One of the main questions asked to gauge producers brand loyalty is responding to the statement of "I consider myself loyal to the brands of expendable items I buy" on a Likert scale with the answer 1 indicating strongly disagree, 2 indicating disagree, 3 indicating neither disagree nor agree, 4 indicating agree, and 5 indicating strongly agree. These responses were further grouped into two categories with the first category containing 5 – strongly agree and 4 – agree and the second category containing 1 – strongly disagree, 2 – disagree, and 3 – neither disagree nor agree. Around 39% of the respondents answered that they agreed or strongly agreed with this statement. A binomial logistic model is used to estimate and test the hypothesized relationships. The two categories, represented as a binary variable, are the dependent variable in this model with explanatory variables of characteristics of the farm and farmer, farmer beliefs and attitudes, and the importance of product characteristics. Three demographic variables have resulting statistically significant negative coefficients. These variables are total annual farm sales in dollars (SALES), a binary variable that equals 1 if the farmer's age is between 35 and 54 (AGE54), and a binary variable that equals 1 if the operation produces cotton (COTTON). Five other variables have proven to be statistically significant, an index variable ranging from 0 to 1 for reported media exposure (MEDINDEX), a binary variable equaling 1 if food/security regulations are important when making input purchase decisions and 0 otherwise (FOOD), a binary variable equaling 1 if it was reported that branded expendable products offer a higher level of performance and 0 otherwise (PERFORM1), a binary variable equaling 1 if respondent has placed an order for agricultural

inputs online and 0 otherwise (ORDONLINE), and a binary variable equaling 1 if the respondent buys the lowest priced expendable products and 0 otherwise(LOWPRICE1). The variables MEDINDEX, FOOD, and PERFORM all had positive coefficients whereas ORDONLINE and LOWPRICE1 had negative coefficients.

Lai, Olynk Widmar, Gunderson, Widmar, and Ortega (2018) studied agricultural producers' preferences for five key management factors of success: managing output prices, managing production, controlling costs, managing land/equipment/facilities, and managing people. The data for this research was obtained through the 2013 edition of the Large Commercial Producer Survey by the Center for Food and Agricultural Business at Purdue University. Pairwise comparison was used instead of Likert scale or ranking to force participants to choose the importance of one factor over another. This research led to four class segments. Each of the classes had a unique prioritization of the five key management factors that separated the producers. Class 1 was generalized to be production focused. Class 2 was focused on managing people and controlling costs. Class 3 was also focused on people but differed from class 2 in that they were focused on production. And lastly, class 4 was focused on commodity marketing.

Reimer, Downey, and Akridge (2009) studied the market segmentation practices among cooperatives and independently owned crop input retailers while addressing the gaps between Best's seven-step market segmentation framework and retailer practices.

The sample of respondents was 55% or 11 respondents that were agricultural cooperatives, 40% or 8 respondents that were privately owned, independent retailers, and 5% or 1 respondent that was a publicly owned retailer.

The first step in Best's seven-step market segmentation framework is needs-based segmentation. This step groups customers into segments based on similar needs of the customer for solving a specific consumption problem. The second step is segment identification. This step is used to determine the factors that make the segments identifiable. The third step is the assessment of segment attractiveness. This step uses the predetermined segments to assess the attractiveness of each segment. The fourth step is evaluating the profitability of the segments by determining each segment's profitability and net marketing contribution. The fifth step is segment positioning. This step creates a strategy to maximize the value for each segment based on their unique needs and characteristics. The sixth step is using the acid test on each segment to test the attractiveness of the value proposition strategies from step five. Step 7 is marketing-mix strategy.

In this step, a marketing mix is developed and targeted at individual segments. Reimer et al. (2009) takes a unique approach to needs-based segmentation that can be used to create additional value.

Roucan-Kane, Alexander, Boehlje, Downey, and Gray (2011) identified four segments of Convenience, Price, Performance, and Balance for commercial producers' buying behaviors using cluster analysis. The data for this research was obtained from the 2008 edition of the Large Commercial Producer Project conducted by the Center for Food and Agricultural Business at Purdue University.

When asked the percent importance of factors including convenience/location, customer service, price, performance, and support service the Balance segment was the largest representing 59% of farms. Buyers in this segment consider all factors to be around equally important. The Price segment was the second-largest segment representing 18% of farms. Buyers in this segment placed large importance on the price factor with an average weight of 47% followed by customer service being the second most important factor closely followed by price. The Convenience segment accounted for 12% of producers. This segment placed large importance on the convenience/location factor with an average weight of 48% followed by convenience/location. The performance segment also accounted for 12% of producers. This segment placed large importance on the performance factor with an average weight of 50% followed by price.

The demographics of each segment also differed on average. Producers in the balance segment were marginally less educated, tended to be older, and a majority considered themselves primarily crop producers. Producers in the convenience segment were the least educated and oldest as compared to the other segments. They were also the most likely to primarily have a livestock operation. Producers in the price segment were the second most educated, the youngest as compared to the other segment, and had the second lowest proportion of primarily livestock operations. Producers in the performance segment were the most educated, were the second youngest, and were the least likely to primarily be a livestock operation.

The authors also compared brand loyalty of the different segments. Consistent with Harbor, Margin, and Akridge (2008), most producers considered themselves to be brand loyal for capital items. However, the level of loyalty differed between the segments. Balance buyers were the most likely to consider themselves brand loyal while price buyers were the least likely to consider themselves brand loyal. Next, the correlation between considering oneself brand loyal and socioeconomic characteristics such as age and level of education. There was not a significant

correlation between brand loyalty and socioeconomic characteristics, further indicating that market segmentation should be based on buying behaviors and not socioeconomic or demographic characteristics.

Borchers et al. (2012) studied how large commercial livestock and crop producers chose their input suppliers by segmenting them according to their buying behaviors. The data used for this research is from the 2008 edition of the Large Commercial Producer Survey conducted by the Center for Food and Agricultural Business at Purdue University. A cluster analysis was conducted for each of the four expendable products focused on: seed, crop protection chemicals, animal health, and animal feed. For livestock producers, they found three buying behavior segments: balance, price, and performance. For crop producers, they found four buying behavior segments: balance, price, performance, and convenience. For all four expendable products, the balance behavioral segment is the largest. They found that producers in the buyer segment consider all input criteria to be of equal importance. However, the relative weights for the input criteria are dependent on the product.

In the seed balance segment, the most important relative factor is product performance with price and customer service tied for the second most important relative factor. In the crop protection chemicals balance segment, the most important relative factor is also product performance with price as the second most important relative factor. On the animal health products and feed balance segments, the most important relative factor is convenience/location with customer service as the second most important relative factor.

The price segment is the second largest for crop protection chemicals, third largest segment for seed, animal health and animal feed. Within the price segment, seed purchasers rank customer service, performance, and convenience/location about equally, crop protection chemical purchasers rank product performance as the second most important factor, livestock input purchasers rank convenience/location as the second most important factor for price buyers.

The product performance segment is the second-largest segment for seed, animal health products and feed and is the third-largest segment for crop protection chemicals. For all expendable products, performance buyers ranked price as the second most important factor.

The convenience segment is only present for the crop expendable products, seed, and crop protection chemicals. This was the smallest segment for crop farmers. In this segment

convenience/location are the most important factor with customer service/information being the second most important factor.

Borchers et al. (2012) also looked at the correlation between these segments and demographic characteristics. They found that for all products performance buyers tended to operate larger farms while crop input convenience buyers operate smaller farms. They also found a higher percentage of college graduates in the price and performance segments as consistent with Gloy and Akridge (1999) and Alexander, Wilson, and Foley (2005).

Brand preferences and loyalty were examined using questions with a 5-point Likert scale with 1 representing strongly disagree and 5 representing strongly agree. It was found that brand loyalty differed by segment along with the product purchased. The price segment buyers tended to be the least likely to be loyal to specific brands for expendable products, seed, chemicals, and animal health. Performance buyers for seed and chemicals tended to be less loyal to specific brands than balance and convenience buyers.

Timberlake (2012) studied the structural changes in the agricultural sector, the importance of the salesperson and producer relationship, and how the producer's way of acquiring inputs has changed. He analyzed survey data from twelve large producers and eighteen retail employees to identify the gaps in the characteristics of a salesperson, activities of a salesperson, drivers of producers' seed decisions, and drivers of producers' crop protection chemical decisions. This research found that as agricultural producers grow and expand, they put more value on price and product performance and less value on services that retailers offer. While Timberlake provided important information, the study was limited by a small sample of survey respondents and limited demographic information.

Terho et al. (2015) analyzed the impact of three dimensions of sales strategies on selling performance. They found that only segmentation directly impacts selling performance while the other two dimensions of prioritization and selling models had an indirect impact. When salespeople can successfully segment their customers in terms of buying behaviors and needs they will optimize their resources and provide increased value positively impacting selling performance.

While these studies analyzed market segmentation for agricultural producers, they did not compare the segmentation of retailers and producers to better understand retailers' understanding of large commercial producers' value proposition.

2.4 Clustering and Multinomial Logistic Regression Model

The following research contributed to offering ways to determine the appropriate number of clusters and successful modeling techniques to segment the agricultural market.

Ketchen and Shook (1996) studied the application of cluster analysis used in strategic management. The concerns about using cluster analysis that the authors elaborated on are: "extensive reliance on researcher judgement that is inherent in cluster analysis" and that perception from critics are "that most applications of cluster analysis in strategy have lacked an underlying theoretical rationale." The selection of the appropriate clustering algorithm is extremely critical for the application of cluster analysis. The two types of clustering algorithms include hierarchical and nonhierarchical. Hierarchical clustering algorithms go through a series of steps which mimic a tree-like structure adding (agglomerative) and deleting (divisive) elements from the clusters. There are five popular agglomerative algorithms of single linkage, complete linkage, average linkage, centroid method, and Ward's method (Hair, Black, Babin, and Anderson, 2019). Each of these algorithms differs in their mathematical procedures, calculating the distance between clusters as well as tendencies in the way they group observations. For example, the average linkage algorithm is biased towards clusters with equal variance, the centroid method is biased towards irregularly shaped clusters and can only be used with interval or ratio data, and lastly, Ward's method is biased towards creating clusters with equivalent numbers of observations and is easily distorted by outliers (The CLUSTER Procedure, 2017).

Hierarchical algorithms are criticized for several problems. First, the underlying structure of the same data is not often known by the researchers, which then creates difficulty in choosing the best algorithm for the data. Secondly, since the algorithm only makes one pass through the data, cluster assignments cannot be modified leading to flawed cluster assignments. And lastly, the stability of the solution is altered when cases are dropped. Nonhierarchical algorithms, also known as k-means algorithms, splits the sample data into a prespecified number of clusters and rearranges the observations in the clusters based on the distance of the observation to the cluster centroid until no observations change clusters. The nonhierarchical clustering algorithm has two potential advantages over the hierarchical clustering algorithm. By allowing observations to switch clusters in the nonhierarchical algorithm, outliers do not have as much of an impact compared to the hierarchical clustering algorithm. Additionally, the multiple passes through the data,

rearranging observations, leads to higher homogeneity within the clusters and heterogeneity between clusters (Aldenderfer & Blashfield, 2011).

The solution recommended to alleviate these problems is to use a two-stage procedure in which a hierarchical algorithm is used to find the correct number of clusters for the dataset, which will then be the starting point/seed values for the nonhierarchical clustering. Research from Milligan (1980) has shown that this method increases the validity of the solution given by the model. (Milligan, 1980) Determining the number of clusters can be done by several techniques including visually inspecting a dendrogram, using an agglomeration coefficient, using cubic clustering criterion (CCC), and using a priori theory. Using multiple of these techniques can help overcome the negatives for each method.

Gloy and Akridge (1999), Alexander, Wilson, and Foley (2005), and Roucan-Kane, Alexander, Boehlje, and Gray (2011) use Ward's hierarchical clustering to determine the optimal number of clusters and seed values for the next step in the process, the k-means algorithm. The k-means algorithm rearranges the observations until no observation change clusters deeming the clusters optimal.

Gloy and Akridge (1999) evaluated several measures to determine the correct number of clusters for the data including cubic clustering criterion, pseudo-F statistic, pseudo-T^2 statistic, and Ward's method. The cubic clustering criterion did not produce a viable solution for this data. The pseudo-F statistic indicated that the data contained three or four clusters. The pseudo-T^2 statistic indicated that the data contained four clusters. However, the pseudo-T^2 statistic indicated some cluster number variation between samples. Ward's method indicated that the data contained four clusters. Ward's method was determined to be the most sensible option for determining the number of clusters for the data because cluster properties did not vary a considerable amount between samples.

Alexander, Wilson, and Foley (2005) used principal and factor analysis were used to find the proper variables for segmentation by identifying highly correlated or redundant variables. After using Ward's method to determine the number of clusters and k-means nonhierarchical algorithm to determine the cluster classifications, multinomial logit regression analysis was used to predict segment membership.

Harbor, Martin, and Akridge (2008) used a binomial logistic model to estimate and test the hypothesized relationships between the dependent and explanatory variables.

Lai, Olynk Widmar, Gunderson, Widmar, and Ortega (2018) used three models to predict the importance of five key management strategies: controlling costs, managing land/equipment/facilities, managing people, managing production, and marketing/price. A multinomial logit regression model was used as a starting point for generating a base model that had the assumption of homogeneity among producers. The latent class model revealed a higher level of importance on one farm management area relative to the other factors.

CHAPTER 3. METHODS

3.1 Overview

This chapter lays out and explains the methods chosen to best answer the hypotheses and problem statement for this study. Because this study involves multiple datasets, multiple analysis methods will be used to extract insights about market segmentation and the gap between how retailers perceive and misclassify their producers.

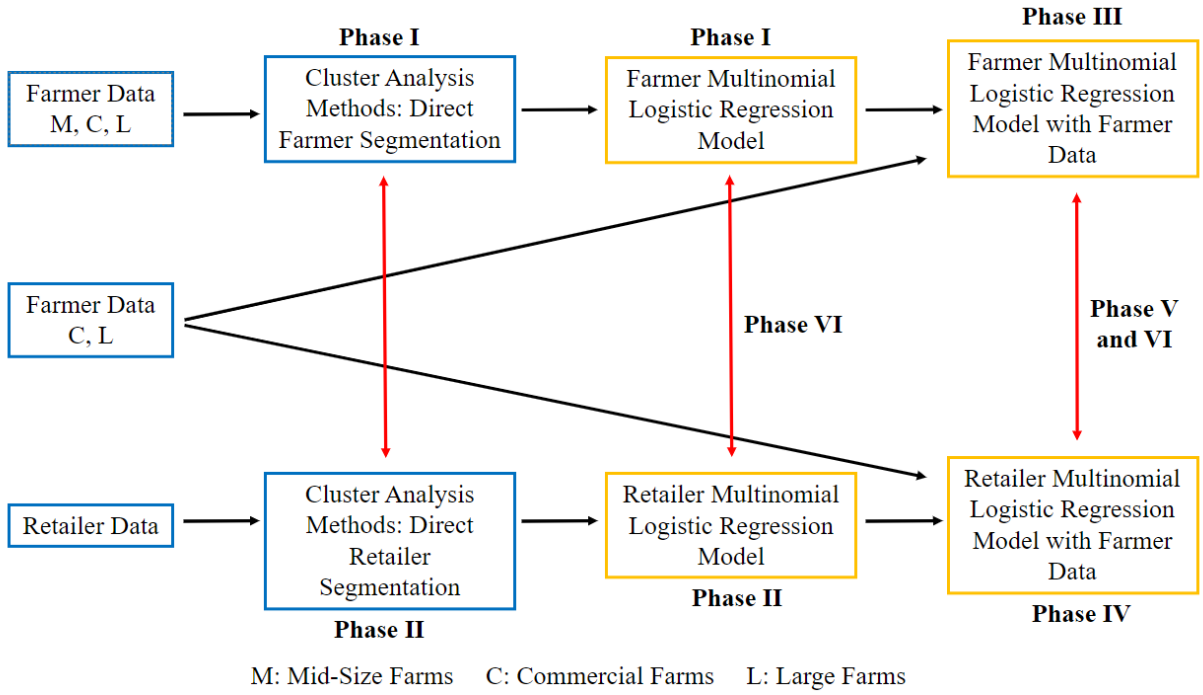

Figure 3-1 Overview of Methods

The methods used in this research will consist of multiple cluster analyses and multinomial logistic regression models in phases as follows:

- Phase I: Farmer Segmentation
    - Farmer data will be used in cluster analysis to create farmer buying segments
        - Uses normalized farmer data containing mid-size, commercial, and large farmers

        - Summary of the proportions of each segment
        - R will be used in this step
    - Farmer data and determined clusters will be used to estimate a producer multinomial logistic regression to extract key defining variables for each cluster
        - Stata will be used in this step
- Phase II: Retailer Segmentation
    - Retailer data will be used in cluster analysis to create retailer-based buying segments
        - Uses normalized retailer data
        - Summary of the proportions of each segment
        - R will be used in this step
    - Retailer data and determined clusters will be used to estimate a multinomial logistic regression to extract key "perceived" defining variables for each cluster
        - Stata will be used in this step
- Phase III: Farmer Segments Identified by Producer Multinomial Logistic Regression Model
    - Uses raw commercial and large farmer data with producer multinomial logistic regression model to predict farmer segmentation
    - These predicted segments will then be used to analyze how each segment on average ranks/responds to the variables in each question. These rankings will be the basis for determining what "type of buyer" each segment represents.
    - Stata will be used in this phase
- Phase IV: Identifying Retailers' Perceptions of Farmer Segments
    - Uses raw commercial and large farmer data with retailer multinomial logistic regression model
    - These predicted segments will then be used to analyze how each segment on average ranks/responds to the variables in each question. These rankings will be compared to the farmer segment rankings in Phase III to determine

which retailer predicted farmer segment is most similar to the "types of buyers" in Phase III
    - Stata will be used in this phase
- Phase V: Confusion Matrix
    - Multi-level confusion matrix with farmer multinomial logistic regression predicted farmer clusters from Phase III as the "actual" or base model and retailer multinomial logistic regression predicted farmer clusters from Phase IV as the "predicted" or challenge model. The results of this matrix will be used to find accuracy measures of retailers' overall prediction as well as individual cluster prediction.
    - R will be used in this phase
- Phase VI: Comparisons and Conclusion
    - Discussion and comparison of the results from the multinomial logistic regression models in Phase I and II as well as the segment proportions determined in Phase III and Phase IV.

Phase I and II are depicted in blue boxes in Figure 2 representing these steps being based on the original data whereas Phase III and IV are depicted in orange boxes representing these steps are estimations. The red arrows represent a point of comparison. The "M", "C", and "L" in the farmer data boxes represent the farm sizes in each dataset with "M" representing mid-size farms, "C" representing commercial size farms, and "L" representing large farms.

3.2 Cluster Analysis Estimation:

To overcome the concern of correlation due to the nature of the ranking questions used to segment the agricultural market, an agglomerative hierarchical clustering method often referred to as the Ward's method will be used to identify the buying segments resulting from each dataset. This method builds a hierarchy of cluster groups from the bottom up by organizing respondents by similar responses to the questions first, then by the closest cluster until the hierarchy is complete, minimizing the total within-cluster variance. During this step will also be when the number of optimal clusters for the data is determined. The number of clusters in this study will be validated using non-parametric statistical tests including the Silhouette Index, Gap Statistic, Duda index, Calinski and Harabasz index, and Pseudo t^2 using the NbClust package in RStudio and Euclidean

Distance as the cluster distance measure. Each of these tests has a different algorithm to determine the optimal number of clusters for this data (Charrad, Ghazzali, Boiteau, & Niknafs, 2014). Formula 3-1 explains how the Euclidean Distance is calculated.

$$Euclidean\ Distance: d(x,y) = \left(\sum_{j=1}^{d}(x_j - y_j)^2\right)^{\frac{1}{2}} \tag{3-1}$$

The maximum value of the Silhouette Index will be used to determine the optimal number of clusters for the data. Formulas 3-2 through 3-6 explain how the Silhouette Index is calculated.

$$Silhouette = \frac{\sum_{i=1}^{n} S(i)}{n},\ Silhouette \in [-1,1], \tag{3-2}$$

Where,

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i); b(i)\}}, \tag{3-3}$$

$$a(i) = \sum_{j\in\{C_r/i\}_{i\,j}^{d}} / n_r - 1, \tag{3-4}$$

$$b(i) = \min\{d_{iC_s}\},\ s \neq r, \tag{3-5}$$

$$d_{iC_s} = \sum_{j\in C_s}{}_{i\,j}^{d} / n_s, \tag{3-6}$$

where a(i) represents the average dissimilarity of the i^{th} object to the objects in cluster C_r. d_{iCs} represents the average dissimilarity of the ith object to all other objects in cluster C_s. (Charrad, Ghazzali, Boiteau, & Niknafs, 2012)

Gap Statistic Index finds the optimal number of clusters by determining the smallest value of q that still satisfies the constraint. Equations 3-7 through 3-12 explain how the Gap Statistic Index is calculated.

$$Gap(q) = \frac{1}{B}\sum_{b=1}^{B} \log W_{qb} - \log W_q, \tag{3-7}$$

$$Gap(q) \geq Gap(q+1) - s_{q+1},\ (q = 1, \ldots, n-2), \tag{3-8}$$

Where,

$$s_q = sd_{q\sqrt{1+1/B}}, \tag{3-9}$$

sd_q is the standard deviation of:

$$\{\log W_{qb}\},\ b = 1, \dots, B\colon\ sd_q = \sqrt{\frac{1}{B}\sum_{b=1}^{B}\left(\log W_{qb} - l\right)^2}, \tag{3-10}$$

$$\bar{l} = \tfrac{1}{B}\textstyle\sum_{b=1}^{B} \log W_{qb}, \tag{3-11}$$

$$W_q = \textstyle\sum_{k=1}^{q}\sum_{i\epsilon C_k}(x_i - \bar{x}_k)(x_i - \bar{x}_k)^T, \tag{3-12}$$

$$q \epsilon (1, \dots, n-2),$$

where W_q is the within-group dispersion matrix for the data clustered into q clusters (Charrad, Ghazzali, Boiteau, & Niknafs, 2012).

The Duda Index finds the optimal number of clusters by imposing a ratio criterion where Je(2) and Je(1) are the sums of squared errors within the clusters when the data is partitioned into two clusters and one cluster. Equations 3-13 through 3-17 explain how the Duda Index is calculated,

$$Duda\ = \frac{J_e(2)}{J_e(1)} = \frac{W_k + W_l}{W_m}, \tag{3-13}$$

$$Duda\ \geq 1 - \frac{2}{\pi p} - z\sqrt{\frac{2\left(1-\frac{8}{\pi^2 p}\right)}{n_m p}} = critValue_{Duda}, \tag{3-14}$$

z = standard normal score,

P = number of variables in the data,

$$W_k = \textstyle\sum_{k=1}^{q}\sum_{i\epsilon C_k}(x_i - c_k)(x_i - c_k)^T, \tag{3-15}$$

$$W_l = \textstyle\sum_{l=1}^{q}\sum_{i\epsilon C_l}(x_i - c_l)(x_i - c_l)^T, \tag{3-16}$$

$$W_m = \textstyle\sum_{m=1}^{q}\sum_{i\epsilon C_m}(x_i - c_m)(x_i - c_m)^T, \tag{3-17}$$

where W represents the within-group dispersion matrix for the data clustered into q clusters for cluster c_i, i = k, l, m. It is assumed that c_m is comprised of the merging of c_k and c_l. (Charrad, Ghazzali, Boiteau, & Niknafs, 2012)

The Calinski and Harabasz (CH) Index finds the value of q in which maximizes CH(q). The value of q represents the optimal number of clusters for the dataset. Equations 3-18 through 3-20 explain how the Calinski and Harabasz Index is calculated,

$$CH(q) = \frac{trace(B_q)/(q-1)}{trace(W_q)/(n-q)}, \tag{3-18}$$

$$W_m = \sum_{m=1}^{q} \sum_{i \epsilon C_m} (x_i - c_m)(x_i - c_m)^T, \tag{3-19}$$

$$B_q = \sum_{k=1}^{q} n_k (c_k - \overline{x})(c_k - \overline{x})^T, \tag{3-20}$$

q = number of clusters,

n = number of observations,

$\overline{x}$ = centroid of data matrix X,

c_k = centroid of cluster C_k,

where W_q represents the within-group dispersion matrix for the data clustered into q clusters. B_q represents the between-group dispersion matrix for the data clustered into q clusters. (Charrad, Ghazzali, Boiteau, & Niknafs, 2012)

The Pseudo t^2 Index finds the smallest number of clusters while keeping the index less than or equal to the critical value. Equation 3-21 through 3-26 explain how the Pseudo t^2 Index is calculated.

$$Pseudo\ t^2 = \frac{V_{kl}}{\frac{W_k + W_l}{n_k + n_l - 2}}, \tag{3-21}$$

Where,

$$V_{kl} = W_m - W_k - W_l, \text{ if } C_m = C_k \cup C_l, \tag{3-22}$$

$$Pseudo\ t^2 \leq \left(\frac{1 - critValue_{Duda}}{critValue_{Duda}}\right) \cdot (n_k + n_l - 2), \tag{3-23}$$

n_k = number of objects in cluster C_k,

n_l = number of objects in cluster C_l,

$$W_k = \sum_{k=1}^{q} \sum_{i \epsilon C_k} (x_i - c_k)(x_i - c_k)^T, \quad (3\text{-}24)$$

$$W_l = \sum_{l=1}^{q} \sum_{i \epsilon C_l} (x_i - c_l)(x_i - c_l)^T, \quad (3\text{-}25)$$

$$W_m = \sum_{m=1}^{q} \sum_{i \epsilon C_m} (x_i - c_m)(x_i - c_m)^T, \quad (3\text{-}26)$$

(Charrad, Ghazzali, Boiteau, & Niknafs, 2012).

The tests above, along with past literature, will be considered to find the optimal number of clusters for the data. While the tests will be biased towards having fewer clusters, keeping a larger number of segments will offer differentiation and interpretability above what only two clusters could offer.

The second step will be a non-hierarchical k-means clustering algorithm that uses the results from the Ward's method as a starting position. A two-step process will be more accurate than just performing the Ward's method on its own. The non-hierarchical k-means algorithm will pass through the data rearranging observations until no observations change, while the Ward's method will only pass through the data one time.

Standardizing the data during the Ward's and k-means steps will be essential due to the questions having different scales. In this research a min-max, also known as 0-1 or linear scaling, normalization will be used to adjust all the observations to be between 0 and 1 for each question. Equation 3-27 explains how to calculate the min-max normalization.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}. \quad (3\text{-}27)$$

(Charrad, Ghazzali, Boiteau, & Niknafs, 2014).

The simplicity of this normalization and dendrogram visualization, along with small standard deviations in terms of the coefficient of variation, points towards this method being optimal over other standardization methods (Almaliki, 2018). The coefficient of variation is the measure of how spread out the values are relative to the variables mean. This ratio determines the size of the standard deviation. A coefficient of variation value less than 1 is considered low and reflects compact patterns of the data. Equations 3-28 and 3-29 explain how to calculate the standard error and coefficient of variation.

$$\sigma = \frac{\sqrt{\Sigma(X_i - X_{bar})^2}}{(n-1)} \tag{3-28}$$

(Donnelly & Abdel-Faouf, 2016)

$$Coefficient\ of\ Variation = \frac{\sigma}{X_{bar}} * 100 \tag{3-29}$$

(Coefficient of Variation: Definition, Formula, Interpretation, Examples & FAQs, n.d.)

3.3 Estimate Multinomial Logistic Regression Models:

Next, multinomial logistic regressions using the farmer and retailer cluster classifications will be used to identify the market segments drivers and statistical significance of the questions used during Phase I and II of this research. Using cluster analysis and multinomial logistic models to segment the agricultural market has proven to be an effective method by Alexander, Wilson, and Foley, 2005 and Lai, Olynk Widmar, Gunderson, Widmar, and Ortega, 2018. A multinomial logistic regression is very similar to a standard logistic regression model in that the model predicts the probability of being classified into a group relative to another. While standard logistic regression models have binary dependent variables, multinomial logistic regressions have multiple classes with a constant base outcome throughout. With 4 clusters and the base outcome being cluster 4, the probabilities are as follows:

$$\Pr(y = 1) = \frac{e^{X\beta^{(1)}}}{e^{X\beta^{(1)}} + e^{X\beta^{(2)}} + e^{X\beta^{(3)}} + 1}, \tag{3-30}$$

$$\Pr(y = 2) = \frac{e^{X\beta^{(2)}}}{e^{X\beta^{(1)}} + e^{X\beta^{(2)}} + e^{X\beta^{(3)}} + 1}, \tag{3-31}$$

$$\Pr(y = 3) = \frac{e^{X\beta^{(3)}}}{e^{X\beta^{(1)}} + e^{X\beta^{(2)}} + e^{X\beta^{(3)}} + 1}, \tag{3-32}$$

$$\Pr(y = 4) = \frac{1}{e^{X\beta^{(1)}} + e^{X\beta^{(2)}} + e^{X\beta^{(3)}} + 1}, \tag{3-33}$$

(mlogit - Multinomial (polytomous) logistic regression)

The first multinomial logistic model uses the farmer data containing mid-size, large, and commercial farms with the farmer cluster classifications assigned after the non-hierarchical clustering algorithm as the dependent variable and the raw responses from the farmer survey segmentation questions used as the independent variables. The results of this model will show the statistical significance, direction, and magnitude of the variables predictability.

The next multinomial logistic model uses the retailer data with the retailer cluster classifications assigned after the non-hierarchical clustering algorithm as the dependent variable and the raw responses from the retailer survey segmentation questions used as the independent variables. The results of this model will show the statistical significance, direction, and magnitude of the variables predictability.

Due to the same questions being used during the cluster analysis step as well as the independent variables in the multinomial logistic models for both the retailer and farmer models with the cluster classification found in the cluster analysis step as the dependent variable there will a problem of over specification in the multinomial logistic models. To prevent having to remove variables and consequently losing valuable information about the retailers' perception and farmers' decision-making behaviors, bootstrapping the standard errors during the multinomial logistic models will be used with 10,000 iterations. Bootstrapping is a type of statistical resampling that uses Monte Carlo methods by drawing a sequence of random samples, also known as bootstraps, from the data to increase the accuracy of the standard error. These samples are then run through the specified model, in this case the multinomial logistic models with the segmentation questions as the independent variables and the cluster classification from the cluster analysis step as the dependent variable. The results from the bootstraps provide an estimation of the sampling distribution (Ratick & Schwarz, 2009).

3.4 Confusion Matrix

Lastly, to understand how accurately retailers can predict the segmentation of their producers, accuracy measures from a multi-level confusion matrix with farmer multinomial logistic regression model predicted farmer clusters from Phase III as the "actual" or base model and predicted farmer clusters from the retailer multinomial logistic regression in Phase IV as the "predicted" or challenge model will be compared. A confusion matrix is a way to visualize the performance of the challenge model against the base model similar to a contingency table. This

comparison will show which of the farmer segments the retailers on average can most accurately predict and which they cannot. (Mohajon, 2020)

3.5 Hypotheses

These methods will be the foundation for answering the hypotheses for this research. Hypothesis 1 will be answered through the comparison of multinomial logistic regression models estimated during Phase I and Phase II as well as in the confusion matrix and accuracy measure in Phase V. Hypothesis 2 will be answered during Phase VI when comparing the retailer predicted farmer cluster proportions with the farmer predicted farmer cluster proportions.

CHAPTER 4. DATA COLLECTION

The data for this research is comprised of producer responses from the Large Commercial Producer Survey and retailer responses from the Agricultural Retailer Survey. These surveys ask the same questions but from a different point of view. The Agricultural Retailer Survey asks how they believe their producers would answer and/or rank the survey questions using the 80-20, meaning the 20% of producers that provide 80% of their business, and the Large Commercial Producer Survey asks how they, the producers, would answer and/or rank the questions in the survey. The data from both surveys was collected between January 2021 and April 2021.

The Large Commercial Producer Survey conducted by the Center for Food and Agricultural Business at Purdue University sampled agricultural producers nationwide. This survey is focused on understanding producers' strategies for success, buying behaviors, risk perceptions around ecommerce and technology and how the retailers perceive the producers' answers to these questions. The survey also has a vast array of demographic questions to help analyze similarities between producers. Stratus Ag Research served as a partner for the Center for Food and Agricultural Business by providing a sample survey format, executing the survey, and collecting the responses online.

The Agricultural Retailer Survey, also conducted by the Center for Food and Agricultural Business at Purdue University, sampled employees from multiple large agricultural retailers. This sample is comprised of employees from companies that work directly with the Center for Food and Agricultural Business and a random sample of agribusiness professionals that work for additional retailers.

The respondent targets for the retailer survey were employees working at agricultural retailers and cooperatives that provide agricultural inputs and services to producers in the United States. The questions that helped answer this study's research question, the segmentation questions, had a total of 555 surveys completed out of 993 surveys started for a question focused completion rate of around 55.9%.

The responses from the farmer survey were targeted as producers of the retailers surveyed. There are two datasets that will be analyzed in this study:

1. Producer data containing mid-size, commercial and large farms
2. Producer data containing only commercial and large farms

There was a total of 1691 surveys started, and of those, 1297 completed the segmentation questions for a completion rate of around 77% for the first dataset containing mid-size, commercial, and large farms. For the dataset containing only commercial and large farms, there were a total of 1030 surveys started, and of those, 758 completed the segmentation questions for a completion rate of around 74%.

The main questions of interest capture important factors and buying behaviors of the farmer through the eyes of the retailer and well as from the farmer. Many of these questions ask retailers and producers to rank the answer options rather than just choosing the most important. In these questions, producers are asked for their opinions and retailers are asked for their opinion on how they think the farmers they serve would answer the questions. Table 4.1 explains the eight segmentation questions and answer options used from the Large Commercial Producer Survey and Agricultural Retailer Survey as well as the reference name for the questions and answers that will be used throughout.

Table 4-1 Segmentation Questions

Farmer Questions and Variables:	**Retailer Questions and Variables:**	**Reference Name:**
Please rank the top three most important factors when selecting which dealer/retailer you purchase products from. (1 is the top choice, then 2 and 3)	**Rank the following factors in order of most important to producers in the area when they select a dealer or retailer to purchase from? Drag and drop the options in the list below so that "the most important" is on top "the lease important" is on the bottom.**	**Dealer/Retailer Attributes Question**
Quality of service performed	Quality of service performed	Service Quality
Availability of service (application, delivery, etc.)	Availability of service (application, delivery, etc.)	Availability of Service
Relationship with salesperson there	Relationship with salesperson	Relationship With Salesperson
Availability of multiple product brands	Availability of multiple brands	Multiple Brands
Past experience with the supplier	Past experience with business	Past Experience
Product returns and warranties	Product returns and warranties	Product Returns and Warranties
Location of the retailer facilities		
Best value or price		

Table 4-1 Continued

In general, which of the following factors is most important in your selection of what products to use? Please rank from 1 to 6, where 1 is "the most important" and 6 is "the least important".	**In general, which of the following factors do you think is most important to large producers in your area in selecting which products to use? Drag and drop the options in the list below so that "the most important" is on top and "the least important" is on the bottom.**	**Product Selection Factor Question**
Availability of product	Availability of product	Availability
Brand of product	Brand of product	Brand
Ease handling of product	Ease handling of product	Ease of Handling
Price of product	Price of product	Price of Product
Effectiveness of product	Effectiveness of product	Effectiveness
Other non-price qualities of product	Other non-price qualities of product	Non-Price Qualities
In general, how would you rank the following characteristics of the salespeople you choose to work with? Please rank from 1 to 4, where 1 is "the most important" and 4 is "the least important".	**In general, what is your perception of how large producers in your area rank the following characteristics of the salespeople they choose to work with? Drag and drop the options in the list below so that "the most important" is on top and "the least important" is on the bottom.**	**Salespeople Characteristics Question**
Relationship - you can trust them with confidential information	Relationship - they can trust the salesperson with confidential information	Relationship
Reliability - you can trust the salesperson will follow through	Reliability - they can trust the salesperson will follow through	Reliability
Credibility - you can trust the information the salesperson provides	Credibility - they can trust the information the salesperson provides	Credibility
Caring - they care about my success as much as their own	Caring - the salesperson cares as much about the customers success as their own	Caring

Table 4-1 Continued

Thinking about the last time you bought inputs from a dealer/retailer, which of the following types of information was most important to you during that purchase?	**Thinking about the last time your large farmer bought inputs from you, rank the types of information you think were most important to him/her during that purchase Drag and drop the options in the list below so that "the most important" is on top and "the least important" is on the bottom.**	**Information Importance Question**
Options for how to save money	Options for how to save money	Information Save Money
Information about different product options	Information about different product options	Information Product Options
Advice about how to maximize profits	Advice about how to maximize profits	Information Maximize Profits
Please rank these attributes in order of importance to you for crop protection	**Please rank these attributes: "Price", "Product Performance", and "Dealer/Retailer Relationship" in order of importance to large producers in your area, with 1 being the "most important attribute" and 3 being "the least important attribute" for crop protection.**	**Crop Protection Attributes Question**
Price	Price	CP Price
Product Performance	Product Performance	CP Performance
Supplier Relationship	Relationship with dealer/retailer	CP Relationship
Using a scale of 1 to 9, where 1 means "strongly disagree" and 9 means "strongly agree", how do you feel about the following?	**Using a scale of "Very Strongly Disagree" to "Very Strongly Agree", indicate your level of agreement with the following statements:**	**Likert Scale Agreement Question**
Significant differences exist in the quality of services between different dealers and retailers	Large producers believe that significant differences exist in the quality of services between different dealers/retailers	Retailer Differences
My relationship with sales people is more important than the relationship I have with the company they represent	Large producers believe that their relationship with the salespeople is more important than the relationship they have with the company they represent	Salesperson Company Relationship
Significant differences exist between generic expendable products (seed, feed/nutrition, etc.) and branded products	large producers believe that significant differences exist between generic expendable products (seed, feed/nutrition, etc.) and branded products	Product Differences
I know more about my expendable products than my dealer or retailer	Large producers believe that they know more about the products they purchase than the dealer/retailer	Know More
Significant differences exist in the quality of information I receive from different dealers and retailers	Large producers believe that significant differences exist in the quality of information they receive from different dealers/retailers	Information Differences

Table 4-1 Continued

For the brands from which you primarily purchase crop protection, please indicate if you agree with the following statements:	**Thinking about the product brands that your large farmer customers primarily purchase, please indicate if you agree with the following statements for crop protection:**	**Brand Loyalty Question**
I consider myself loyal to my primary brand	I believe my large customers are loyal to the brands they purchase	Loyal to Brand
I have recommended this brand to someone else	I believe my large customers recommend the brands they purchase to others	Recommend Brand
I would switch from this brand for a 5% discount	I believe my large customers would switch brands for a 5% discount	Would Switch Brands for a 5% Discount
I would switch from this brand for a 10% discount	I believe my large customers would switch brands for a 10% discount	Would Switch Brands for a 10% Discount
For the dealer/retailer from which you primarily purchase the products listed below, please indicate if you agree with the following statements:	**Thinking about the company you represent, please indicate if you agree with the following statements about your large farmer customers' in relation to crop protection.**	**Company Loyalty Question**
I consider myself loyal to the primary dealer/retailer for crop protection	I believe my large customers are loyal to my company	Loyal to Company
I have recommended this dealer/retailer for crop protection to someone else	I believe my large customers recommend my company to others	Recommend Company
I would switch from this dealer/retailer for a 5% discount in crop protection	I believe my large customers would switch from my company for a 5% discount	Would Switch Companies for a 5% Discount
I would switch from this dealer/retailer for a 10% discount in crop protection	I believe my large customers would switch from my company for a 10% discount	Would Switch Companies for a 10% Discount

The frequencies of each question-and-answer option for both surveys are located in Appendix B in Table B-1 through Table B-35 for the farmer survey and Table B-36 through B-70 for the retailer survey. To make the survey questions comparable, modification of the results was essential.

In the Dealer/Retailer Attributes Question, the number of rankings and quantity of answer options differed between the farmer and retailer survey. The farmer survey asks producers to rank the top three choices and offers additional answer options of location of the retailer facilities and

best value or price whereas the retailer survey asks for a ranking of all 6 variables. To make these responses comparable, the variables given ranks 4, 5, or 6 in the retailer survey were designated as "4". In the farmer survey results only the variables consistent in both surveys were analyzed. Out of these variables, if not given a rank in the top 3, a rank of 4 was designated.

The farm size cutoffs were determined by the United States Department of Agriculture's Economic Research Service and were as follows: Mid-Size farms had gross farm income (GCFI) greater than $149,999 but less than or equal to $999,999 annually, Commercial farms had GCFI greater than $999,999 but less than or equal to $4,999,999 annually, and lastly, Large farms had GCFI greater than $4,999,999 annually. Table 4-2 summarizes the average responses to each of the segmentation questions and demographics for all farms, mid-size, commercial, and large farms as well as the average response from the retailers to the segmentation questions. The demographics chosen to analyze are farm size in acres, farm revenue in USD, and age in years. Table 4-3 summarizes the proportion of each category of education in each farm size. (Hoppe & MacDoland, 2013)

Table 4-2 Demographic and Segmentation Question Averages By Farm Size

Variable	**All Mid-Size, Commercial, and Large Farms n=1297** $\bar{x}$	**Mid-Size Farms n=544** $\bar{x}$	**Commercial Farms n=599** $\bar{x}$	**Large Farms n=154** $\bar{x}$	**Retailer n=555** $\bar{x}$
Size (acres)	5094 [b]	1550 [a]	4205 [b]	21067 [a, b]	-
Revenue ($)	$6,985,084 [a]	$3,388,927 [b]	$7,948,344 [a]	$15,900,000 [c]	-
Age (Years)	56 [b]	58 [c]	55 [a]	55 [a, b]	-
Education (Percent of Farm Size Category)					
High School or Less	21%	24%	17%	15%	-
Less than Bachelor's Degree	23%	24%	23%	19%	-
4 Year Degree	55%	49%	57%	63%	
Graduate Degree	3%	3%	3%	2%	-

Table 4-2 Continued

Dealer/Retailer Attribute Question (ranked 1 to 3, lower number is higher ranked)					
Service Quality	3.06 [a]	3.03 [a]	3.07 [a]	3.12 [a]	2.27 [b]
Availability of Service	3.33 [a]	3.29 [a]	3.36 [a]	3.34 [a]	1.82 [b]
Relationship With Salesperson	3.23 [a]	3.27 [a]	3.19 [a]	3.25 [a]	2.88 [b]
Multiple Brands	3.70 [a]	3.66 [a]	3.73 [a]	3.69 [a]	3.85 [b]
Past Experience	3.14 [b, c]	3.16 [a, b]	3.17 [a, c]	2.99 [b]	3.24 [a]
Product Returns and Warranties	3.81 [a, b]	3.84 [a]	3.79 [a, b]	3.79 [b]	3.94 [c]
Product Selection Factor Question (ranked 1 to 5, lower number is higher ranked)					
Availability	3.29 [a]	3.28 [a]	3.27 [a]	3.42 [a]	3.69 [b]
Brand	4.71 [a]	4.70 [a]	4.72 [a]	4.71 [a]	3.68 [a]
Ease of Handling	3.89 [b]	4.00 [a]	3.81 [b]	3.84 [a, b, c]	3.69 [c]
Price of Product	2.25 [a, b]	2.19 [a]	2.29 [b]	2.28 [a, b]	2.23 [a, b]
Effectiveness	1.64 [a]	1.64 [a]	1.63 [a]	1.66 [a]	1.65 [a]
Non-Price Qualities	5.22 [a, b]	5.20 [a, b]	5.28 [a]	5.08 [b, c]	5.06 [c]
Salespeople Characteristics Question (ranked 1 to 4, lower number of higher ranked)					
Relationship	2.71 [a]	2.76 [a]	2.72 [a]	2.5 [b]	2.47 [b]
Reliability	2.21 [a, b]	2.17 [a]	2.22 [b]	2.36 [c]	2.24 [a, b, c]
Credibility	2.05 [a]	2.03 [a]	2.06 [a]	2.06 [a, b]	2.21 [b]
Caring	3.02 [a]	3.04 [a]	3.00 [a]	3.08 [a]	3.08 [a]
Information Importance Question (ranked 1 to 3, lower number is higher ranked)					
Information Save Money	1.99 [a, b]	2.05 [b]	1.96 [a]	1.90 [a]	2.18 [c]
Information Product Options	2.18 [a]	2.14 [a]	2.18 [a]	2.27 [a]	2.16 [a]
Information Maximize Profits	1.84 [a]	1.81 [a]	1.86 [a]	1.84 [a]	1.66 [b]
Crop Protection Attributes Question (ranked 1 to 3, lower number is higher ranked)					
CP Price	1.91 [a, b]	1.96 [a]	1.88 [b]	1.87 [a, b]	1.96 [a]
CP Performance	1.54 [a, b]	1.49 [a]	1.58 [b]	1.58 [a, b]	1.71 [c]
CP Relationship	2.54 [a]	2.55 [a]	2.54 [a]	2.55 [a]	2.33 [b]

Table 4-2 Continued

Likert Scale Agreement Question (Likert 1 to 9 with 2 being disagree to 9 being agree)					
Retailer Differences	6.96 [a]	6.89 [a]	6.95 [a, b]	7.2 [b]	6.26 [c]
Salesperson Company Relationship	7.03 [a, b]	7.06 [a, b]	6.95 [a, b]	7.21 [a]	6.88 [b]
Product Differences	5.35 [b]	5.59 [c]	5.16 [a]	5.23 [a, b]	5.03 [a]
Know More	4.85 [a]	4.67 [b]	4.95 [a]	5.12 [a]	4.35 [c]
Information Differences	6.39 [a]	6.39 [a]	6.36 [a]	6.53 [a]	6.32 [a]
Brand Loyalty Question (yes/no questions where 1 is yes and 0 is no)					
Loyal to Brand	0.49 [a, b]	0.51 [a, b]	0.46 [a]	0.53 [b]	0.42 [c]
Recommend Brand	0.46 [a, b]	0.48 [a, b]	0.44 [a]	0.53 [b]	0.62 [c]
Would Switch Brands for a 5% Discount	0.41 [a]	0.38 [a]	0.41 [a]	0.49 [b]	0.62 [c]
Would Switch Brands for a 10% Discount	0.62 [a]	0.61 [a]	0.63 [a]	0.64 [a]	0.87 [b]
Company Loyalty Question (yes/no questions where 1 is yes and 0 is no)					
Loyal to Company	0.6 [a]	0.61 [a]	0.59 [a]	0.62 [a]	0.57 [b]
Recommend company	0.52 [a]	0.5 [a]	0.52 [a]	0.6 [b]	0.64 [c]
Would Switch Companies for a 5% Discount	0.41 [a]	0.4 [a]	0.41 [a]	0.44 [a]	0.44 [b]
Would Switch Companies for a 10% Discount	0.6 [a]	0.6 [a]	0.58 [a]	0.64 [a]	0.77 [b]

Note: A matching letter indicates that the means for these variables in the same row are not statistically different.

The demographics from Table 4-2 shows that revenue is directly related with farm size, as when farm size increases the average farm revenue also increases. However, the increase of average farm acres compared to farm revenue are not proportional. The segmentation question averages offered initial insights into the differences between each farm sizes buying behaviors. While the average numbers for each question and variable themselves are significant, understanding how the average respondent would rank the order of or give higher importance to the variable options offers additional conclusions. When comparing all mid-size, commercial, and large farms to the retailer column for the segmentation questions, many of the questions have similar or the same ranking order for the answer options. However, the magnitude of the averages

differs. In the next chapter, farmers and retailers will be segmented using cluster analysis to group the most likeminded farmers and the most likeminded retailers according to the segmentation questions around buying behaviors. These clusters will be the base for the multinomial logistic regression models resulting in identifying farmer segments and retailers' perceptions of farmer segments. Finally, the accuracy at which retailers can predict farmers into the correct segment will be explored.

CHAPTER 5. ANALYSIS AND RESULTS

5.1 Introduction

This chapter reports the results from this research in the order proposed during chapter 3, with five phases. Phase I contains cluster analysis and a multinomial logistic regression model using farmer segmentation question data focused on producers' buying behaviors with mid-size, commercial, and large farm size. Phase II contains cluster analysis and a multinomial logistic regression model using retailer segmentation question data focused on how they perceive their producers' buying behaviors. Phase III contains the identification of farmer segments using the producer multinomial regression model estimated in Phase I and raw farmer data containing only commercial and large farm size observations. The cluster names, representing the type of buyer each segment embodies, were determined in this phase as well. Phase IV contains the identification of retailers' perceptions of farmer segments using the retailer multinomial logistic regression model estimated in Phase II and raw farmer data containing only commercial and large farm size observations. Similar to Phase III, cluster names, representing the type of buyer each segment perceives was determined and matched by similarity of ranking/variables importance to the existing farmer segment names in Phase IV. Phase V reports the accuracy of retailers' ability to correctly predict their producers into the correct segment using a multilevel confusion matrix. Phase VI summarizes the results and offers a comparison between the results from each phase.

5.2 Phase I: Farmer

In section 5.2, cluster analysis using min-max normalized farmer data containing mid-size, commercial, and large farm size observations determined the segments that were then used as the dependent variable in the producer multinomial logistic regression model with the farmers answers to the segmentation questions used as the independent variables. This model will continue to be used in Phase III and results from Phase III in Phase V.

5.2.1 Cluster Analysis

In section 5.2.1, cluster analysis using the Ward's method with non-parametric statistical tests along with previous literature determined and validated the optimal number of clusters for

this data to be 4 clusters. Next a non-hierarchical k-means clustering algorithm used the number of clusters determined from the Ward's method to segment the farmer data into four clusters. Table 5-1 summarizes the frequency and proportions of each segment,

Table 5-1 Farmer Clusters

	Freq.	Percent
1	177	13.65
2	370	28.53
3	370	28.53
4	380	29.30
Total	1,297	100

where the fourth cluster contains the highest frequency with 380 observations, representing 29.30% of the data. The second and third cluster were tied for the next highest observations with 370 observations, representing 28.53% of the data each. The first cluster has the lowest frequency with 177 observations, representing 13.65% of the data. These segments are the dependent variable in the producer multinomial logistic regression model in the next step in section 5.2.2.

5.2.2 Multinomial Logistic Regression Model

In section 5.2.2, a multinomial logistic regression model was estimated using the farmer cluster classifications determined in section 5.2.1 as the dependent variable and raw farmer segmentation question data containing mid-size, commercial, and large farm size observations as the independent variables to extract the statistically significant variables that determined the farmer cluster classifications. In this model the standard errors were bootstrapped with 10,000 iterations and had the base outcome of cluster 4. Tables 5-2 through 5-4 summarize the statistically significant variables for each cluster relative cluster 4 from the producer multinomial logistic regression model. The complete results from this model can be found in Appendix B.

Table 5-2 Variables Statistically Significant in Predicting Cluster 1 Compared to Cluster 4

Variable	Parameter and Statistical Significance
CP Price	-51.72***
CP Performance	-44.47***
Retailer Differences	-9.580**
Product Differences	-10.32**
Loyal to Brand	-179.2***
Recommend Brand	-38.03*
Would Switch Brands for a 5% Discount	150.7***
Would Switch Brands for a 10% Discount	87.05***
Loyal to Company	-200.9***
Recommend Company	-123.8***
Would Switch Companies for a 5% Discount	189.6***
Would Switch Companies for a 10% Discount	130.8***
Information Save Money	-33.18***

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5-2 summarizes the variables statistically significant for predicting cluster 1 relative to cluster 4 from the producer multinomial logistic regression model. Variables Retailer Differences, Product Differences, Loyal to Brand, Recommend Brand, Loyal to Company, and Recommend Company had a negative direction of impact. Because a higher value for these variables indicates higher agreement or importance, this negative direction of impact indicated that when these variables increase in importance, the respondent is less likely to be placed in cluster 1 relative to cluster 4. Variables CP Price, CP Performance, and Information Save Money also had a negative direction of impact. However, since a lower value for these variables indicates higher importance, when these variables decrease in importance, the respondent is less likely to be placed in cluster 1 relative to cluster 4. Variables Would Switch Brands for a 5% Discount, Would Switch Brands for a 10% Discount, Would Switch Companies for a 5% Discount, and Would Switch Companies for a 10% Discount had a positive direction of impact. Because a higher value for these variables indicates higher agreement, this positive direction of impact indicated that when these variables increase in importance, the respondent is more likely to be placed in cluster 1 relative to cluster 4.

Table 5-3 Variables Statistically Significant in Predicting Cluster 2 Compared to Cluster 4

Variable	Parameter and Statistical Significance
Relationship With Salesperson	33.60***
Past Experience	20.33**
Salesperson Company Relationship	18.01**
Recommend Brand	93.11***
Would Switch Brands for a 5% Discount	155.7***
Would Switch Brands for a 10% Discount	158.8***
Would Switch Companies for a 5% Discount	184.8***
Would Switch Companies for a 10% Discount	157.2***

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5-3 summarizes the variables statistically significant for predicting cluster 2 relative to cluster 4 from the producer multinomial logistic regression model. Variables Salesperson Company Relationship, Recommend Brand, Would Switch Brands for a 5% Discount, Would Switch Brands for a 10% Discount, Would Switch Companies for a 5% Discount, and Would Switch Companies for a 10% Discount had a positive direction of impact. Because a higher value for these variables indicates higher importance, when these variables increase in importance, the respondent is more likely to be placed in cluster 2 relative to cluster 4. Variables Relationship With Salesperson and Past Experience also had a positive direction of impact. However, since a lower value indicates higher importance, this positive direction of impact indicated that when these variables decrease in importance, the respondent is more likely to be place in cluster 2 relative to cluster 4.

Table 5-4 Variables Statistically Significant in Predicting Cluster 3 Compared to Cluster 4

Variables	Parameter and Statistical Significance
Service Quality	6.622*
Relationship With Salesperson	7.664*
Multiple Brands	-10.69*
CP Price	-47.58***
CP Performance	-34.96***
Loyal to Brand	-125.9***
Recommend Brand	-54.75***
Would Switch Brands for a 5% Discount	45.00*
Would Switch Brands for a 10% Discount	125.0***
Loyal to Company	-166.7***
Recommend Company	-125.7***
Would Switch Companies for a 10% Discount	168.4***
Reliability	-8.841*
Information Product Options	12.10**

**** p<0.01, ** p<0.05, * p<0.1*

Table 5.4 summarizes the variables statistically significant for predicting cluster 3 relative to cluster 4 from the producer multinomial logistic regression model. Variables Loyal to Brand, Recommend Brand, Loyal to Company and Recommend Company had a negative direction of impact. Because a higher value for these variables indicates higher agreement, this negative direction of impact indicated that when these variables increase in importance, the respondent is less likely to be placed in cluster 3 relative to cluster 4. Variables Multiple Brands, CP Price, CP Performance, and Reliability also had a negative direction of impact. However, since a lower value for these variables indicates higher importance, when these variables decrease in importance, the respondent is less likely to be placed in cluster 3 relative to cluster 4. Variables Would Switch Brands for 5% Discount, Would Switch Brands for a 10% Discount, and Would Switch Companies for a 10% Discount had a positive direction of impact. Because a higher value for these variables indicates higher agreement, this positive direction of impact indicated that when these variables increase in importance, the respondent is more likely to be placed in cluster 3 relative to cluster 4. Variables Service Quality, Relationship With Salesperson, Multiple Brands, and Information Product Options also had a positive direction of impact. However, since a lower value for these variables indicates higher importance, when these variables decrease in importance, the respondent is more likely to be placed in cluster 3 relative to cluster 4.

Table 5-5 Variables Dropped Due to Multicollinearity

Variables
CP Relationship
Caring
Information Maximize Profits
Non-Price Qualities

Table 5-5 summarizes the variables that were omitted during the producer multinomial logistic regression model due to multicollinearity. Out of 10,000 bootstrap iterations 3,841 replications converged.

5.3 Phase II: Retailer

In section 5.3, cluster analysis using min-max normalized retailer data determined the segments that were then used as the dependent variable in the retailer multinomial logistic regression model with retailers' answers to the segmentation questions used as the independent variables. This model will continue to be used in Phase IV and results from Phase IV in Phase V.

5.3.1 Cluster Analysis

In section 5.3.1, cluster analysis using the Ward's method with non-parametric statistical tests along with previous literature determined and validated the optimal number of clusters for this data to also be 4 clusters. Next, similar to the farmer data, a non-hierarchical k-means clustering algorithm used the number of clusters determined from the Ward's method to segment the retailer data into four clusters. Table 5.6 summarizes the frequency and proportions of each segment,

Table 5-6 Retailer Clusters

	Freq.	Percent
1	179	32.25
2	151	27.21
3	114	20.54
4	111	20.00
Total	555	100

where the first cluster contains the highest frequency with 179 observations, representing 32.25% of the data. The second cluster had the next highest frequency with 151, representing 27.21% of the data, followed by cluster three with 114 observations, representing 20.54% of the data. The fourth cluster had the lowest frequency with 111 observations, representing 20% of the data. These segments are the dependent variable in the retailer multinomial logistic regression model in the next step in section 5.3.2.

5.3.2 Multinomial Logistic Regression Model

In section 5.3.2, a multinomial logistic regression model was estimated using the retailer cluster classifications determined in section 5.3.1 as the dependent variable and raw retailer segmentation question data as the independent variables to extract the statistically significant variables that determined the retailer cluster classifications. In this model, similar to producer multinomial logistic regression model, the standard errors were bootstrapped with 10,000 iterations and had the base outcome of cluster 4. Tables 5-7 through 5-9 summarize the statistically significant variables for each cluster relative to cluster 4 from the retailer multinomial logistic regression model. The complete results from this model can be found in Appendix B.

Table 5-7 Variables Statistically Significant in Predicting Cluster 1 Compared to Cluster 4

Variable	Parameter and Statistical Significance
CP Price	-133.4***
CP Performance	-58.70***
Loyal to Brand	-62.17***
Recommend Brand	-58.71***
Would Switch Brands for a 5% Discount	128.8***
Would Switch Brands for a 10% Discount	82.61**
Loyal to Company	-146.8***
Recommend Company	-144.6***
Would Switch Companies for a 5% Discount	131.5***
Would Switch Companies for a 10% Discount	175.4***
Information Save Money	-63.18***
Price of Product	-21.26*

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5-7 summarizes the variables statistically significant for predicting cluster 1 relative to cluster 4 for the retailer multinomial logistic regression model. Variables Loyal to Brand, Recommend Brand, Loyal to Company, and Recommend Company had a negative direction of impact. Because a higher value for these variables indicates higher agreement or importance, this negative direction of impact indicated that when these variables increase in importance, the respondent is less likely to be placed in cluster 1 relative to cluster 4. Variables CP Price, CP Performance, Information Save Money, and Price of Product also had a negative direction of impact. However, since a lower value for these variables indicates higher importance, when these variables decrease in importance, the respondent is less likely to be places in cluster 1 relative to cluster 4. Variables Would Switch Brands for a 5% Discount, Would Switch Brands for a 10% Discount, Would Switch Companies for a 5% Discount, and Would Switch Companies for a 10% Discount had a positive direction of impact. Because a higher value for these variables indicates higher agreement, this positive direction of impact indicated that when these variables increase in importance, the respondent is more likely to be placed in cluster 1 relative to cluster 4.

The retailer model was edited and rerun after determining which retailer clusters were most similar to each farmer cluster so that the cluster numbers match to create the highest comparison and interpretability between the farmer clusters predicted by the producer multinomial logistic

regression model in Phase III and the farmer clusters predicted by the retailer multinomial logistic regression model in Phase IV.

Table 5-8 Variables Statistically Significant in Predicting Cluster 2 Compared to Cluster 4

Variable	Parameter and Statistical Significance
CP Price	-87.46***
CP Performance	-62.95***
Would Switch Brands for a 5% Discount	148.1***
Would Switch Brands for a 10% Discount	40.94*
Loyal to Company	65.10**
Would Switch Companies for a 5% Discount	98.72***
Would Switch Companies for a 10% Discount	160.0***
Reliability	-19.20**
Credibility	-20.73**
Information Save Money	-21.98*
Effectiveness	-25.33**

**** p<0.01, ** p<0.05, * p<0.1*

Table 5-8 summarizes the variables statistically significant for predicting cluster 2 relative to cluster 4 from the retailer multinomial logistic regression model. Variables CP Price, CP Performance, Reliability, Credibility, Information Save Money, and Effectiveness had a negative direction of impact. Because a lower value for these variables indicates higher importance, when these variables decrease in importance, the respondent is less likely to be placed in cluster 2 relative to cluster 4. Variables Would Switch Brands for a 5% Discount, Would Switch Brands for a 10% Discount, Loyal to Company, Would Switch Companies for a 5% Discount, and Would Switch Companies for a 10% Discount had a positive direction of impact. Because a higher value for these variables indicates higher agreement, this positive direction of impact indicated that when these variables increase in importance, the respondent is more likely to be placed in cluster 2 relative to cluster 4.

Table 5-9 Variables Statistically Significant in Predicting Cluster 3 Compared to Cluster 4

Variable	Parameter and Statistical Significance
CP Performance	37.14**
Would Switch Brands for a 10% Discount	65.24***
Recommend Company	-58.47**
Would Switch Companies for a 10% Discount	152.4***

**** p<0.01, ** p<0.05, * p<0.1*

Table 5-9 summarizes the variables statistically significant for predicting cluster 3 relative to cluster 4 from the retailer multinomial logistic regression model. Recommend Company had a negative direction of impact. Because a higher value for this variable indicates higher agreement, this negative direction of impact indicated that when this variable increases in importance, the respondent is less likely to be placed in cluster 3 relative to cluster 4. Variables Would Switch Brands for a 10% Discount and Would Switch Companies for a 10% Discount had a positive direction of impact. Because a higher value for these variables indicates higher agreement, this positive direction of impact indicated that when these variables increase in importance, the respondent is more likely to be placed in cluster 3 relative to cluster 4. CP Performance also had a positive direction of impact. However, since a lower value for this variable indicates higher importance, when the importance for this variable decrease, the respondent is more likely to be placed in cluster 3 relative to cluster 4.

Table 5-10 Variables Dropped Due to Multicollinearity

Variables
Product Returns and Warranties
CP Relationship
Caring
Information Maximize Profits
Non-Price Qualities

Table 5-10 summarizes the variables that were omitted during the retailer multinomial logistic regression model due to multicollinearity. Out of 10,000 bootstrap iterations 411 replications converged.

5.4 Phase III: Identifying Farmer Segments Using the Producer Multinomial Logistic Regression Model

Section 5.4 summarizes how farmer segments were identified using the producer multinomial logistic regression model resulting in cluster names determining what type of buyer each segment embodies.

First, the clusters for raw farmer data containing commercial and large farm size observations were estimated using the producer multinomial logistic regression model with a scalar threshold of 0.1. These classifications were then used to determine the averages for each segmentation question/variable, which gave insight into how the average respondent in each cluster would rank or give importance to each question and variable. Table 5-11 summarizes the farmers average responses predicted by the producer multinomial logistic regression model to each of the segmentation questions and demographics for all commercial and large farms as well as by cluster.

Table 5-11 Farmer Predicted Farmer Demographic and Segmentation Question Averages by Cluster

Variable	**All Large and Commercial Farms n=758** $\bar{x}$	**Farmer Predicted Farmer Cluster 1 n=110** $\bar{x}$	**Farmer Predicted Farmer Cluster 2 n=221** $\bar{x}$	**Farmer Predicted Farmer Cluster 3 n=208** $\bar{x}$	**Farmer Predicted Farmer Cluster 4 n=219** $\bar{x}$
% of Producers	100% [a, b, c]	15% [a, b, c]	29% [a]	27% [b]	29% [c]
Average Size (acres)*	5,244 [e]	4,539 [a]	4,797 [b]	5,589 [c]	5,755 [d]
Median Size (acres)*	3,600 [e]	3,860 [a]	3,713 [b]	3,500 [c]	3,600 [d]
Revenue ($1000's)	$9,561 [c]	$8,490 [a]	$10,353 [a, b]	$11,082 [d]	$7,905 [b]
Age (Years)	55 [a]	50 [a, b, c]	57 [a]	55 [b]	55 [c]

Table 5-11 Continued

Education (Percent of Segment)					
High School or Less	17% [c]	8% [a, b, c]	15% [a]	16% [b]	24% [a, b, c]
Less than Bachelor's Degree	22% [c]	12% [a, b, c]	28% [a, b, c]	20% [a]	22% [b]
4 Year Degree	58% [a, b, c]	80% [a, b]	51% [a]	63% [a, c]	53% [b, c]
Graduate Degree	3% [a, b]	0% [a, b]	6% [a]	1% [a]	1% [b]
Dealer/Retailer Attribute Question (ranked 1 to 4, lower number is higher ranked)					
Service Quality	3.08 [d]	3.26 [a, b, d]	2.95 [a]	3.12 [c]	3.07 [b]
Availability of Service	3.36 [d]	3.46 [a]	3.42 [b]	3.33 [c]	3.27 [a]
Relationship With Salesperson	3.20 [a, c]	3.42 [a, b]	3.01 [a, c]	3.42 [c, d]	3.08 [b, d]
Multiple Brands	3.72 [e]	3.01 [a, b]	3.82 [c, d, e]	3.66 [a, c]	3.64 [b, d]
Past Experience	3.13 [d]	3.32 [a, b, d]	3.08 [a]	3.19 [c]	3.03 [b]
Product Returns and Warranties	3.79 [e]	3.73 [a]	3.81 [b]	3.78 [c]	3.80 [d]
Product Selection Factor Question (ranked 1 to 5, lower number is higher ranked)					
Availability	3.31 [c]	3.36 [a]	3.56 [b, c]	3.20 [b]	3.13 [a, c]
Brand	4.72 [c]	4.90 [a]	4.79 [b]	4.88 [c]	4.40 [a, b, c]
Ease of Handling	3.82 [e]	3.83 [a]	3.90 [b]	3.81 [c]	3.73 [d]
Price of Product	2.29 [c, d]	1.96 [a, d]	2.33 [a, b]	2.14 [b, c]	2.54 [a, c]
Effectiveness	1.64 [d]	1.75 [a]	1.53 [a, b]	1.58 [c]	1.74 [b]
Non-Price Qualities	5.24 [b, c]	5.20 [a]	4.89 [a, b]	5.39 [b]	5.47 [a, c]
Salespeople Characteristics Question (ranked 1 to 4, lower number is higher ranked)					
Relationship	2.68 [d]	2.73 [a]	2.78 [b]	2.72 [c]	2.51 [a, b, c, d]
Reliability	2.24 [d]	2.15 [a]	2.27 [c]	2.16 [b]	2.34 [a, b]
Credibility	2.06 [e]	2.03 [a]	2.08 [b]	2.04 [c]	2.08 [d]
Caring	3.02 [d]	3.10 [a]	2.88 [a, b, c, d]	3.08 [b]	3.07 [c]

Table 5-11 Continued

Information Importance Question (ranked 1 to 3, lower number is higher ranked)					
Information Save Money	1.95 [a]	1.77 [a, b]	2.06 [a, c]	1.87 [c, d]	1.99 [b, d]
Information Product Options	2.20 [a]	2.23 [c]	2.08 [a]	2.38 [a, b]	2.13 [b]
Information Maximize Profits	1.86 [a]	2.00 [a]	1.86 [c]	1.75 [a, b]	1.88 [b]
Crop Protection Attributes Question (ranked 1 to 3, lower number is higher ranked)					
CP Price	1.88 [c, d]	1.65 [a, d]	1.90 [a, b]	1.75 [b, c]	2.11 [a, c]
CP Performance	1.58 [e]	1.60 [a]	1.56 [b]	1.53 [c]	1.64 [d]
CP Relationship	2.54 [c, d]	2.75 [a, d]	2.55 [a, b]	2.72 [b, c]	2.26 [a, c]
Likert Scale Agreement Question (Likert 1 to 9 with 1 being disagree to 9 being agree)					
Retailer Differences	7.01 [d]	6.45 [a, b, d]	6.93 [a, c]	7.06 [b, c]	7.31 [a, d]
Salesperson Company Relationship	6.98 [d]	6.45 [a, b, c, d]	7.12 [a]	6.90 [b]	7.19 [c]
Product Differences	5.18 [c]	4.19 [a, b, c]	5.20 [a]	5.27 [b]	5.56 [a, c]
Know More	4.99 [b]	5.37 [a, b]	4.90 [a]	5.18 [c]	4.71 [b, c]
Information Differences	6.40 [d]	6.15 [a]	6.28 [b]	6.49 [c]	6.56 [a]
Brand Loyalty Question (yes/no questions where 1 is yes and 0 is no)					
Loyal to Brand	0.48 [a]	0.02 [a]	0.95 [a]	0.10 [a]	0.58 [a]
Recommend Brand	0.46 [b]	0.16 [a, b]	0.99 [a, b]	0.09 [a, b]	0.42 [a]
Would Switch Brands for a 5% Discount	0.43 [a]	0.87 [a]	0.94 [a]	0.03 [a]	0.07 [a]
Would Switch Brands for a 10% Discount	0.63 [a]	0.49 [a]	1.00 [a]	0.74 [a]	0.23 [a]

Table 5-11 Continued

Company Loyalty Question (yes/no questions where 1 is yes and 0 is no)					
Loyal to Company	0.60 [a]	0.05 [a]	0.97 [a]	0.20 [a]	0.87 [a]
Recommend Company	0.54 [a, c]	0.16 [a]	0.99 [a, b]	0.13 [b, c]	0.66 [a, c]
Would Switch Companies for a 5% Discount	0.42 [a, d]	0.94 [a, b]	0.93 [c, d]	0.02 [a, c]	0.02 [b, d]
Would Switch Companies for a 10% Discount	0.60 [b]	0.53 [a]	1.00 [a, b]	0.77 [a, b]	0.06 [a, b]

Note: A matching letter indicates that the means for these variables in the same row are statistically different using T Tests for the mean comparison, Mood's Median Tests for the median comparison, and Two Proportion Z-Tests for proportion comparisons.

** One outlier variable dropped for farm size*

The averages for the segmentation questions as well as the ranking of these averages in each segmentation question were the basis for determining what type of buyer each cluster represented.

After analyzing these rankings, it was determined that,

- Cluster 1 represented Economic buyers,
- Cluster 2 represented Agronomic buyers,
- Cluster 3 represented Business buyers, and
- Cluster 4 represented Relationship buyers.

Cluster 1 was determined to represent Economic buyers because these respondents were very likely to switch brands and companies for a discount, would not call themselves loyal to a brand or company, were unlikely to recommend a brand or company to someone else, and unlike any other cluster, ranked Information Save Money first in the Information Importance Question.

Cluster 2 was determined to represent Agronomic buyers because while these producers may be very likely to switch brands or companies for a 5% or 10% discount, they also think of themselves as brand and company loyal and would be very likely to recommend a brand or company to someone else. Cluster 2 as compared to cluster 3 appears to care more about the salesperson they work with based on the ranking of their responses to the Dealer/Retailer Attribute Question and Likert Scale Agreement Question.

Cluster 3 was determined to represent Business buyers because of these producers' high likelihood of switching brands or companies for a 10% discount and because of importance placed on their experience with the retailer based off their responses to the variable Past Experience in the Dealer/Retailer Attributes Question and the variable Retailer Differences in the Likert Scale Agreement Question.

Cluster 4 was determined to represent Relationship buyers because of their low likelihood of switching brands or companies for a 5% or 10% discount, high likelihood of calling themselves loyal to a brand or company, and high probability of recommending a brand or company to someone else. Table 5-12 summarizes the frequencies and proportions of each type of buyer segment along with acres in each segment.

Table 5-12 Farmer Buyer Segments

Cluster	**Frequency**	**Proportion**	**Total Acres**	**Proportion of Acres**
Economic Buyer	110	14.51%	499,320	9%
Agronomic Buyer	221	29.16%	1,059,386	18%
Business Buyer	208	27.44%	2,957,848	51%
Relationship Buyer	219	28.89%	1,260,364	22%

Next the demographics for each buyer segment from Table 5-11 were analyzed. The Economic buyer segment had the highest percentage of respondents with at least a 4-year degree, followed by the Business buyer segment. The Relationship buyer segment had the highest proportion of respondents with high school or less as their highest degree. These cluster names, segmentation question averages, and ranking of variables in each question will be the foundation for the retailer cluster classifications representation of perceived producers buying behaviors in the next section.

5.5 Phase IV: Identifying Retailers' Perceptions of Farmer Segments

Section 5.5 summarizes how retailer perceived farmer segments were identified using the retailer multinomial logistic regression model resulting in cluster names determining what type of buyer each segment perceives.

First, the clusters for raw farmer data containing commercial and large farm size observations were estimated using the retailer multinomial logistic regression model with a scalar threshold of 0.1. These classifications were then used to determine the averages for each segmentation question/variable, which gave insight into how the average respondent in each cluster would rank or give importance to each question and variable. Table 5-13 summarizes the farmers average responses predicted by the retailer's multinomial logistic regression model to each of the segmentation questions and demographics for all commercial and large farms as well as by cluster.

Table 5-13 Retailer Predicted Demographic and Segmentation Question Averages by Cluster

Variable	**All Large and Commercial Farms n=758** $\bar{x}$	**Retailer Predicted Farmer Cluster 1 n=114** $\bar{x}$	**Retailer Predicted Farmer Cluster 2 n=266** $\bar{x}$	**Retailer Predicted Farmer Cluster 3 n=133** $\bar{x}$	**Retailer Predicted Farmer Cluster 4 n=245** $\bar{x}$
% of Producers	100% [a, b, c, d]	15% [a, b]	35% [a, c]	18% [c, d]	32% [b, d]
Average Size (acres)	5,244 [e]	6,085 [a]	4,861 [b]	4,632 [c]	5,629 [d]
Median Size (acres)	3,600 [e]	3,200 [a]	3,935 [b]	3,525 [c]	3,700 [d]
Revenue ($1000's)	$9,561 [d]	$12,845 [b]	$10,173 [a]	$8,716 [c]	$7,888 [a]
Age (Years)	55 [e]	53 [a]	55 [b]	54 [c]	55 [d]
Education (Percent of Segment)					
High School or Less	17% [a]	13% [a]	14% [b]	14% [c]	24% [a, b, c]
Less than Bachelor's Degree	22% [b]	21% [a]	26% [a, b, c]	18% [b]	21% [c]
4 Year Degree	58% [e]	65% [a, b, e]	55% [a, c]	67% [c, d, f]	55% [b, d]
Graduate Degree	3% [a, b, c]	1% [a]	6% [a, b, c]	1% [b]	1% [c]

Table 5-13 Continued

Dealer/Retailer Attribute Question (ranked 1 to 4, lower number is higher ranked)					
Service Quality	3.08 [d]	3.03 [a]	3.02 [b]	3.28 [a, b, c, d]	3.05 [c]
Availability of Service	3.36 [b]	3.27 [a]	3.56 [a, b, c]	3.13 [b]	3.31 [c]
Relationship With Salesperson	3.20 [d]	3.16 [a]	3.17 [b]	3.41 [a, b, c, d]	3.15 [c]
Multiple Brands	3.72 [a, b, c]	3.53 [a, b]	3.85 [a, c]	3.83 [b, d]	3.62 [c, d]
Past Experience	3.13 [e]	3.22 [a]	3.16 [b]	3.18 [c]	3.03 [d]
Product Returns and Warranties	3.79 [e]	3.77 [a]	3.81 [b]	3.74 [c]	3.8 [d]
Product Selection Factor Question (ranked 1 to 5, lower number is higher ranked)					
Availability	3.31 [d]	3.18 [a]	3.44 [a, b]	3.37 [c]	3.18 [b]
Brand	4.72 [a, b, c]	4.92 [a, b]	4.96 [c, d]	4.47 [a, c]	4.49 [b, d]
Ease of Handling	3.82 [d]	3.84 [b]	3.90 [a]	3.68 [a]	3.79 [c]
Price of Product	2.29 [c]	1.89 [a, b, c]	2.25 [a]	2.35 [b]	2.47 [a, c]
Effectiveness	1.64 [c]	1.90 [a, c]	1.45 [a, b, c]	1.71 [b]	1.67 [a]
Non-Price Qualities	5.24 [b, c]	5.25 [a]	4.99 [a, b, c]	5.41 [b]	5.4 [c]
Salespeople Characteristics Question (ranked 1 to 4, lower number is higher ranked)					
Relationship	2.68 [b]	2.78 [a]	2.85 [b, c]	2.32 [a, b, d]	2.64 [c, d]
Reliability	2.24 [e]	2.14 [a]	2.21 [b]	2.3 [c]	2.29 [d]
Credibility	2.06 [b]	2.10 [a]	1.89 [a, b, c]	2.29 [b]	2.10 [c]
Caring	3.02 [e]	2.98 [a]	3.05 [b]	3.10 [c]	2.97 [d]
Information Importance Question (ranked 1 to 3, lower number is higher ranked)					
Information Save Money	1.95 [c]	1.46 [a, b, c]	2.01 [a]	2.14 [a, b, c]	1.99 [b]
Information Product Options	2.20 [c]	2.61 [a, b, c]	2.06 [a, c]	2.23 [a]	2.14 [b]
Information Maximize Profits	1.86 [d]	1.93 [a]	1.92 [b]	1.63 [a, b, c, d]	1.87 [c]

Table 5-13 Continued

Crop Protection Attributes Question (ranked 1 to 3, lower number is higher ranked)					
CP Price	1.88 [b, d]	1.44 [a, b]	1.84 [a, c]	2.01 [a, d]	2.07 [b, c]
CP Performance	1.58 [b]	1.67 [a]	1.47 [a, b]	1.77 [b, c]	1.56 [c]
CP Relationship	2.54 [a]	2.89 [a]	2.69 [a]	2.23 [a]	2.38 [a]
Likert Scale Agreement Question (Likert 1 to 9 with 1 being disagree to 9 being agree)					
Retailer Differences	7.01 [d]	6.82 [a]	6.81 [b]	7.13 [c]	7.24 [a, b, d]
Salesperson Company Relationship	6.98 [d]	7.06 [b]	6.92 [c]	6.77 [a]	7.13 [a]
Product Differences	5.18 [b]	4.61 [a, b]	4.98 [c, d]	5.35 [a, c]	5.56 [b, d]
Know More	4.99 [c]	5.45 [a, b, c]	4.87 [a]	5.08 [b]	4.87 [b]
Information Differences	6.40 [d]	6.25 [b]	6.23 [a]	6.70 [a, d]	6.49 [c]
Brand Loyalty Question (yes/no questions where 1 is yes and 0 is no)					
Loyal to Brand	0.48 [a]	0.05 [a, b]	0.73 [a, b]	0.38 [a]	0.45 [b]
Recommend Brand	0.46 [a, b]	0.10 [a, b]	0.79 [a, b]	0.32 [a]	0.34 [b]
Would Switch Brands for a 5% Discount	0.43 [c]	0.47 [a, b]	0.92 [a, b, c]	0.10 [a, c]	0.06 [b, c]
Would Switch Brands for a 10% Discount	0.63 [a, b]	0.75 [a]	0.87 [a, b]	0.79 [b]	0.24 [a, b]

Table 5-13 Continued

Company Loyalty Question (yes/no questions where 1 is yes and 0 is no)					
Loyal to Company	0.60 [a]	0.04 [a]	0.82 [a]	0.44 [a]	0.70 [a]
Recommend Company	0.54 [b]	0.06 [a, b]	0.84 [a, b]	0.34 [a, b]	0.54 [a]
Would Switch Companies for a 5% Discount	0.42 [a]	0.54 [a]	0.88 [a]	0.10 [a]	0.04 [a]
Would Switch Companies for a 10% Discount	0.60 [a, b]	0.78 [a, b]	0.89 [a]	0.92 [b]	0.02 [a, b]

Note: A matching letter indicates that the means for these variables in the same row are statistically different using T Tests for the mean comparison, Mood's Median Tests for the median comparison, and Two Proportion Z-Tests for proportion comparisons.

** One outlier variable dropped for farm size*

The averages for the segmentation questions as well as the ranking of these averages in each segmentation question were the basis for determining what type of buyer each cluster represented.

After analyzing these rankings, it was determined that,

- Cluster 1 represented Economic buyers,
- Cluster 2 represented Agronomic buyers,
- Cluster 3 represented Business buyers, and
- Cluster 4 represented Relationship buyers.

Cluster 1 was determined to represent Economic buyers because these respondents unlike any other cluster ranked CP Price first on average in the Crop Protection Attributes Question and Information Save Money first in the Information Importance Question. They also were very likely to switch brands and companies for a discount, would not call themselves loyal to a brand or company, and were unlikely to recommend a brand or company to someone else.

Cluster 2 was determined to represent Agronomic buyers because while these producers may be very likely to switch brands or companies for a 5% discount, they also think of themselves as brand and company loyal and would be very likely to recommend a brand or company to someone else.

Cluster 3 was determined to represent Business buyers because of these producers' high likelihood of switching brands or companies for a 10% discount.

Cluster 4 was determined to represent Relationship buyers because of their low likelihood of switching brands or companies for a 5% or 10% discount, high likelihood of calling themselves loyal to a brand or company, and high probability of recommending a brand or company to someone else. Table 5-14 summarizes the frequencies and proportions of each type of buyer segment.

Table 5-14 Retailer Buyer Segments

Cluster	Frequency	Proportion	Total Acres	Proportion of Acres
Economic Buyer	114	15.04%	2,488,655	43%
Agronomic Buyer	266	35.09%	1,293,004	22%
Business Buyer	133	17.55%	616,037	11%
Relationship Buyer	245	32.32%	1,379,222	24%

Next the demographics for each perceived buyer segment from Table 5-13 were analyzed. The Business buyer segment had the heist percentage of respondents with at least a 4-year degree followed by the Economic buyer segment. The Relationship buyer segment had the highest proportion of respondents with high school or less as their highest degree. These retailer perceived farmer cluster names, segmentation averages, and ranking of variables as well as the producer cluster names, segmentation averages, and ranking of variables from Phase III will be used in the next section to determine retailers' ability to accurately predict producers into the correct segment based off buying behaviors.

5.6 Phase V: Confusion Matrix

A multilevel confusion matrix was used to analyze the performance measures of the farmer clusters predicted by retailer multinomial logistic regression model. These performance measures measure overall accuracy of predictability as well as analyze the clusters accuracy individually

with sensitivity, specificity, precision, error rate, and accuracy. The formulas for these measures are:

Overall Accuracy measures the overall ability of the challenge model; in this case the retailer multinomial logit; to predict the base model; in this case the farmer multinomial logit model. The equation is:

$$Overall\ Accuracy = \frac{TP_1 + TP_2 + TP_3 + TP_4}{Total\ number\ of\ observations} \tag{5-1}$$

Sensitivity measures the proportion of all positive observations from the base model that were correctly predicted positive by the challenge model. The equation is:

$$Sensitivity = \frac{TP_i}{TP_i + FN_i} \tag{5-2}$$

Specificity measures the proportion of all negative observations from the base model that were correctly predicted negative by the challenge model. The equation is:

$$Specificity = \frac{TN_i}{FP_i + TN_i} \tag{5-3}$$

Precision measures the proportion of predictions as a positive class from the challenge model that were actually positive in the base model. The equation is:

$$Precision = \frac{TP_i}{TP_i + FP_i} \tag{5-4}$$

Error Rate measures the proportion of predictions from the challenge model that were incorrect. The equation is

$$Error\ Rate = \frac{FP_i + FN_i}{TP_i + TN_i + FP_i + FN_i} \tag{5-5}$$

Accuracy measures the proportion of predictions from the challenge model that were correct. The equation is:

$$Accuracy = \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \tag{5-6}$$

Where:

TP = True Positive

TN = True Negative

FP = False Positive; Type 1 Error

FN = False Negative; Type 2 Error

(Mohajon, 2020)

Table 5-15 is the contingency table for farmer clusters and farmer predicted clusters from the retailer multinomial logistic regression model representing the "actual" and "predicted" values.

Table 5-15 Confusion Matrix

		Challenge Model				
		Economic	Agronomic	Business	Relationship	Total
Base Model	Economic	**53**	44	3	10	110
	Agronomic	6	**194**	21	0	221
	Business	55	16	**92**	45	208
	Relationship	0	12	17	**190**	219
	Total	114	266	133	245	758

Table 5-16 through 5-19 are the confusion matrices cell classifications for each individual cluster accuracy measures. Table 5-20 summarizes the accuracy measures for retailer's ability to predict farmers into the correct segment.

Table 5-16 Confusion Matrix for Economic Buyers

Base Model \ Challenge Model	Economic	Agronomic	Business	Relationship
Economic	TP	FN	FN	FN
Agronomic	FP	TN	TN	TN
Business	FP	TN	TN	TN
Relationship	FP	TN	TN	TN

Table 5-17 Confusion Matrix for Agronomic Buyers

Base Model \ Challenge Model	Economic	Agronomic	Business	Relationship
Economic	TN	FP	TN	TN
Agronomic	FN	TP	FN	FN
Business	TN	FP	TN	TN
Relationship	TN	FP	TN	TN

Table 5-18 Confusion Matrix for Business Buyers

Base Model \ Challenge Model	Economic	Agronomic	Business	Relationship
Economic	TN	TN	FP	TN
Agronomic	TN	TN	FP	TN
Business	FN	FN	TP	FN
Relationship	TN	TN	FP	TN

Table 5-19 Confusion Matrix for Relationship Buyers

Base Model \ Challenge Model	Economic	Agronomic	Business	Relationship
Economic	TN	TN	TN	FP
Agronomic	TN	TN	TN	FP
Business	TN	TN	TN	FP
Relationship	FN	FN	FN	TP

Table 5-20 Accuracy Measures

Overall Accuracy	
0.7	
Sensitivity	
Economic	0.48
Agronomic	0.88
Business	0.44
Relationship	0.87
Specificity	
Economic	0.91
Agronomic	0.87
Business	0.93
Relationship	0.9
Precision	
Economic	0.46
Agronomic	0.73
Business	0.69
Relationship	0.78

Table 5-20 Continued

Error Rate	
Economic	0.16
Agronomic	0.13
Business	0.21
Relationship	0.11
Accuracy	
Economic	0.84
Agronomic	0.87
Business	0.79
Relationship	0.89

The Overall Accuracy measure of the farmer clusters predicted by the retailer multinomial logistic regression model is 0.70, meaning that 70% of respondents were predicted correctly with a p-value of < 2.2e-16. The accuracy measures for the farmer clusters predicted by the retailer multinomial logistic regression model varied in each individual cluster.

The Agronomic buyer segment had the highest sensitivity compared to the other segments. A sensitivity measure of 0.88 indicated that 88% of the actual Agronomic respondents were predicted into the Agronomic segment by the retailer multinomial logistic regression, also called the true positive rate.

The Business buyer segment had the highest specificity compared to the other segments. A specificity rate of 0.93 indicated that 93% of actual respondents not in the Business buyer segment were predicted to be in segments other than the Business buyer segment, also called the true negative rate.

The Relationship buyer segment had the highest precision, accuracy, and lowest error rate. A precision measure of 0.78 indicated that 78% of respondents predicted to be in the Relationship buyer segment were in the producer Relationship buyer segment, also called the positive predictive value. An accuracy rate of 0.89 indicated that 89% of the values either predicted in the Relationship buyer segment were in the relationship buyer segment or predicted to not be in the Relationship buyer segment and were not in the Relationship buyer segment. An error rate of 0.11 indicated that 11% of values either predicted in the Relationship buyer segment were not in the Relationship

buyer segment or predicted to not be in the Relationship buyer segment but were in the Relationship buyer segment.

The higher the sensitivity, specificity, prediction, accuracy, and lower the error rate the higher the predication accuracy. The retailer multinomial logistic regression model had the highest accurate predictability with the Relationship buyer segment containing the highest precision, highest accuracy, and lowest error rate as compared to the other segments. The Agronomic buyer segment had the next highest accurate predictability with the highest sensitivity, but lowest specificity as compared to the other segments. The Economic buyer and Business buyer segments had the lowest accuracy with the Economic buyer segment having the lowest precision rate and the Business buyer segment having the lowest sensitivity, lowest accuracy, and highest error rate. However, the Business buyer segment did have the highest specificity compared to the other segments.

In the next section, the comparison of statistically significant variables from the producer multinomial logistic regression model or base model estimated in Phase I and the retailer multinomial logistic regression model or challenge model estimated in Phase II will be compared to extract further insights into the retailer biases causing these differences in accuracy measures between segments.

In the next section, further insights into the retailer biases causing these differences in accuracy measures between the segments will be explored. First, the statistically significant variables from the base model estimated in Phase I will be compared to the challenge model in Phase II to understand the gaps between the model drivers. Next, the segment proportions determined after Phase III and Phase IV will be compared and tested to see if they are statistically different. Lastly, the demographics and average responses to the variables in the segmentation questions will be compared within each segment to further extract which variables cause the gap in value proposition between retailers and farmers as well as the differences in accuracy measures between the segments.

5.7 Phase VI: Comparisons

In section 5.7 the results from each phase of this research will be presented and summarized. Comparisons will start with the significant variables from multinomial logistic regression model for the farmer data in Phase I, and for the retailer data in Phase II. Next the proportions and

demographics for the farmer segments predicted by the producer multinomial logistic regression model will be compared to those of the farmer predicted by the retailer multinomial logistic regression model.

5.7.1 Multinomial Logistic Regression Model Comparison

Section 5.7.1 compares the statistically significant variables between the producer multinomial logistic regression model and retailer multinomial logistic regression for each segment compared to the fourth segment, the Relationship buyers. Table 5-21 summarizes the variables statistically significant in predicting the Economic buyer segment, the Agronomic buyer segment, and the Business buyer segment compared to the Relationship buyer segment for the segments using farmer data and the producer multinomial logistic regression model, "farmer segments", and farmer data and the retailer multinomial logistic regression model, "retailer segment".

Table 5-21 Variables Statistically Significant for Each Segment Compared to the Relationship Buyer Segment

	Economic Buyer Segment Compared to Relationship Buyer Segment		Agronomic Buyer Segment Compared to Relationship Buyer Segment		Business Buyer Segment Compared to Relationship Buyer Segment	
	Farmer	**Retailer**	**Farmer**	**Retailer**	**Farmer**	**Retailer**
Dealer/Retailer Attribute Question						
Service Quality	-	-	-	-	6.622*	-
Availability of Service	-	-	-	-	-	-
Relationship With Salesperson	-	-	33.60***	-	7.664*	-
Multiple Brands	-	-	-	-	-10.69*	-
Past Experience	-	-	20.33**	-	-	-
Product Returns and Warranties	-	-	-	-	-	-

Table 5-21 Continued

Product Selection Factor Question						
Availability	-	-	-	-	-	-
Brand	-	-	-	-	-	-
Ease of Handling	-	-	-	-	-	-
Price of Product	-	-21.26*	-	-	-	-
Effectiveness	-	-	-	-25.33**	-	-
Non-Price Qualities	-	-	-	-	-	-
Salespeople Characteristics Question						
Relationship	-	-	-	-	-	-
Reliability	-	-	-	-19.20**	-8.841*	-
Credibility	-	-	-	-20.73**	-	-
Caring	-	-	-	-	-	-
Information Importance Question						
Information Save Money	-	-63.18***	-	-21.98*	-	-
Information Product Options	-	-	-	-	12.10**	-
Information Maximize Profits	-	-	-	-	-	-
Crop Protection Attributes Question						
CP Price	-51.72***	-133.4***	-	-87.46***	-47.58***	37.14**
CP Performance	-44.47***	-58.70***	-	-62.95***	-34.96***	-
CP Relationship	-	-	-	-	-	-
Likert Scale Agreement Question						
Retailer Differences	-9.580**	-	-	-	-	-
Salesperson Company Relationship	-	-	18.01**	-	-	-
Product Differences	-10.32**	-	-	-	-	-
Know More	-	-	-	-	-	-
Information Differences	-	-	-	-	-	-

Table 5-21 Continued

Brand Loyalty Question						
Loyal to Brand	-179.2***	-62.17***	-	-	-125.9***	-
Recommend Brand	-38.03*	-58.71***	93.11***	-	-54.75***	-
Would Switch Brands for a 5% Discount	150.7***	128.8***	155.7***	148.1***	45.00*	-
Would Switch Brands for a 10% Discount	87.05***	82.61**	158.8***	40.94*	125.0***	65.24***
Company Loyalty Question						
Loyal to Company	-200.9***	-146.8***	-	65.10**	-166.7***	-
Recommend company	-123.8***	-144.6***	-	-	-125.7***	-58.47**
Would Switch Companies for a 5% Discount	-	131.5***	184.8***	98.72***	-	-
Would Switch Companies for a 10% Discount	130.8***	175.4***	157.2***	160.0***	168.4***	152.4***

**** p<0.01, ** p<0.05, * p<0.1*

Comparing the farmer and retailer columns for the Economic buyer segment compared to the Relationship buyer segment shows that these models have mostly the same statistically significant variables and direction of impact aside from a few variables. The variables Retailer Differences and Product Differences were statistically significant in the producer multinomial logistic regression model but not in the retailer multinomial logistic regression model. The variable Price of Product was statistically significant in the retailer multinomial logistic regression model but not in the producer multinomial logistic regression model.

Comparing the farmer and retailer columns for the Agronomic buyer segment compared to the Relationship buyer segment shows that these models, while having a few of the same statistically significant variables and direction of impact including Would Switch Brands for a 5% Discount, Would Switch Brands for a 10% Discount, Would Switch Companies for a 5% Discount, and Would Switch Companies for a 10% Discount, these two models have a higher number of variables differing in statistical significance than they do similar variables. The variables Relationship With Salesperson, Past Experience, Salesperson Company Relationship, and Recommend Brand were statistically significant in the producer multinomial logistic regression but not in the retailer multinomial logistic regression model. The variables CP Price, CP

Performance, Loyal to Company, Reliability, Credibility, Information Save Money, and Effectiveness were statistically significant in the retailer multinomial logistic regression model, but not in the producer multinomial logistic regression model.

Comparing the farmer and retailer columns for the Business buyer segment compared to the Relationship buyer segment shows that while having a few of the same statistically significant variables with the same direction of including Would Switch Brands for a 10% Discount, Recommend Company, and Would Switch Companies for a 10% Discount, they differed on the direction of impact on the variable CP Performance as well as statistical significance for many other variables. Variables Service Quality, Relationship With Salesperson, Multiple Brands, CP Price, Loyal to Brand, Recommend Brand, Would Switch Brands for a 5% Discount, Would Switch Brands for a 10% Discount, Loyal to Company, Reliability, and Information Product Options were statistically significant in the producer multinomial logistic regression model but not in the retailer multinomial logistic regression model.

The differences between the variables' statistical significance and direction of impact in these two models defined which variables influenced the gap between the farmer and retailer buying segments.

5.7.2 Farmer Predicted and Retailer Predicted Farmer Cluster Comparison

In section 5.7.2, the farmer predicted farmer cluster proportions and retailer predicted farmer proportions estimated in Phase III and Phase IV will be compared. Table 5-21 has the same proportions as in Table 5-13 and Table 5-16 but merged into one table with the buyer segment names instead of cluster numbers.

Table 5-22 Farmer Predicted Versus Retailer Predicted Farmer Cluster Proportions Comparison

Segment	Farmer Predicted Farmer Cluster Proportion	Retailer Predicted Farmer Cluster Proportion
Economic	14.51% [a]	15.04% [b]
Agronomic	29.16% [c]	35.09% [c]
Business	27.44% [d]	17.55% [d]
Relationship	28.89% [e]	32.32% [f]

Note: A matching letter indicates that the segment proportions between the farmer predicted farmer clusters and the retailer predicted farmer clusters were statistically different using a Two Proportion Z-Test.

From Table 5-22, while comparing not only the proportions of each segment but also the order in which the segments rank in order of size, the ranking of the segment proportions is consistent between the farmer predicted farmer clusters and retailer predicted farmer clusters. However, using a Two Proportion Z-Test determined that the proportion size for the Agronomic buyer and Business buyer segment were statistically different.

5.7.3 Segment Comparison

Table 5-21 compares the demographics and average responses predicted by the producer multinomial logistic regression model or base model and predicted by the retailer multinomial logistic regression model or challenge model for the variables in each of the segmentation questions. These values come from Table 5-11 and Table 5-13.

Table 5-21 Farmer and Retailer Comparison by Segment

Variable	All Large and Commercial Farms n=758	Predicted Economic Buyer Segment		Predicted Agronomic Buyer Segment		Predicted Business Buyer Segment		Predicted Relationship Buyer Segment	
		Farmer n=110	Retailer n=114	Farmer n=221	Retailer n=266	Farmer n=208	Retailer n=133	Farmer n=219	Retailer n=245
% of Producers	100%	15% [c]	15% [d]	29% [a]	35% [a]	27% [b]	18% [b]	29% [e]	32% [f]
Average Size (acres)*	5,244	4,539 [a]	6,085 [b]	4,797 [c]	4,861 [d]	5,589 [e]	4,632 [f]	5,755 [g]	5,629 [h]
Median Size (acres)*	3,600	3,860 [a]	3,200 [b]	3,713 [c]	3,935 [d]	3,500 [e]	3,525 [f]	3,600 [g]	3,700 [h]
Revenue ($1000's)	$9,561	$8,491 [a]	$12,700 [b]	$10,354 [c]	$10,200 [d]	$11,029 [e]	$8,716 [f]	$7,905 [g]	$7,887 [h]
Age (Years)	55	50 [a]	53 [a]	57 [b]	55 [c]	55 [d]	54 [e]	55 [f]	55 [g]

Table 5-21 Continued

Education (Percent of Segment)									
High School or Less	17%	8% [a]	13% [a]	15% [b]	14% [c]	16% [d]	14% [e]	24% [f]	24% [g]
Less than Bachelor's Degree	22%	12% [a]	21% [a]	28% [b]	26% [c]	20% [d]	18% [e]	22% [f]	21% [g]
4 Year Degree	58%	80% [a]	65% [a]	51% [b]	55% [c]	63% [d]	67% [e]	53% [f]	55% [g]
Graduate Degree	3%	0% [a]	1% [a]	6% [b]	6% [c]	1% [d]	1% [e]	1% [f]	1% [g]
Dealer/Retailer Attribute Question (ranked 1 to 4, lower number is higher ranked)									
Service Quality	3.08	3.26 [a]	3.03 [a]	2.95 [b]	3.02 [c]	3.12 [d]	3.28 [e]	3.07 [f]	3.05 [g]
Availability of Service	3.36	3.46 [b]	3.27 [c]	3.42 [d]	3.56 [e]	3.33 [a]	3.13 [a]	3.27 [f]	3.31 [g]
Relationship With Salesperson	3.20	3.42 [a]	3.16 [a]	3.01 [b]	3.17 [c]	3.42 [d]	3.41 [e]	3.08 [f]	3.15 [g]
Multiple Brands	3.72	3.01 [a]	3.53 [a]	3.82 [c]	3.85 [d]	3.66 [b]	3.83 [b]	3.64 [e]	3.62 [f]

Table 5-21 Continued

Past Experience	3.13	3.32 [a]	3.22 [b]	3.08 [c]	3.16 [d]	3.19 [e]	3.18 [f]	3.03 [g]	3.03 [h]
Product Returns and Warranties	3.79	3.73 [a]	3.77 [b]	3.81 [c]	3.81 [d]	3.78 [e]	3.74 [f]	3.80 [g]	3.80 [h]
Product Selection Factor Question (ranked 1 to 5, lower number is higher ranked)									
Availability	3.31	3.36 [a]	3.18 [b]	3.56 [c]	3.44 [d]	3.20 [e]	3.37 [f]	3.13 [g]	3.18 [h]
Brand	4.72	4.90 [b]	4.92 [c]	4.79 [d]	4.96 [e]	4.88 [a]	4.47 [a]	4.40 [f]	4.49 [g]
Ease of Handling	3.82	3.83 [a]	3.84 [b]	3.90 [c]	3.90 [d]	3.81 [e]	3.68 [f]	3.73 [g]	3.79 [h]
Price of Product	2.29	1.96 [b]	1.89 [c]	2.33 [d]	2.25 [e]	2.14 [a]	2.35 [a]	2.54 [f]	2.47 [g]
Effectiveness	1.64	1.75 [a]	1.90 [b]	1.53 [c]	1.45 [d]	1.58 [e]	1.71 [f]	1.74 [g]	1.67 [h]
Non-Price Qualities	5.24	5.20 [a]	5.25 [b]	4.89 [c]	4.99 [d]	5.39 [e]	5.41 [f]	5.47 [g]	5.40 [h]

Table 5-21 Continued

Salespeople Characteristics Question (ranked 1 to 4, lower number is higher ranked)									
Relationship	2.68	2.73 [b]	2.78 [c]	2.78 [d]	2.85 [e]	2.72 [a]	2.32 [a]	2.51 [f]	2.64 [g]
Reliability	2.24	2.15 [a]	2.14 [b]	2.27 [c]	2.21 [d]	2.16 [e]	2.30 [f]	2.34 [g]	2.29 [h]
Credibility	2.06	2.03 [c]	2.10 [d]	2.08 [a]	1.89 [a]	2.04 [b]	2.29 [b]	2.08 [e]	2.10 [f]
Caring	3.02	3.10 [b]	2.98 [c]	2.88 [a]	3.05 [a]	3.08 [d]	3.10 [e]	3.07 [f]	2.97 [g]
Information Importance Question (ranked 1 to 3, lower number is higher ranked)									
Information Save Money	1.95	1.77 [a]	1.46 [a]	2.06 [c]	2.01 [d]	1.87 [b]	2.14 [b]	1.99 [e]	1.99 [f]
Information Product Options	2.20	2.23 [a]	2.61 [a]	2.08 [c]	2.06 [d]	2.38 [b]	2.23 [b]	2.13 [e]	2.14 [f]
Information Maximize Profits	1.86	2.00 [a]	1.93 [b]	1.86 [c]	1.92 [d]	1.75 [e]	1.63 [f]	1.88 [g]	1.87 [h]

Table 5-21 Continued

Crop Protection Attributes Question (ranked 1 to 3, lower number is higher ranked)									
CP Price	1.88	1.65 [a]	1.44 [a]	1.90 [c]	1.84 [d]	1.75 [b]	2.01 [b]	2.11 [e]	2.07 [f]
CP Performance	1.58	1.60 [b]	1.67 [c]	1.56 [d]	1.47 [e]	1.53 [a]	1.77 [a]	1.64 [f]	1.56 [g]
CP Relationship	2.54	2.75 [a]	2.89 [a]	2.55 [b]	2.69 [b]	2.72 [c]	2.23 [c]	2.26 [d]	2.38 [d]
Likert Scale Agreement Question (Likert 1 to 9 with 1 being disagree to 9 being agree)									
Retailer Differences	7.01	6.45 [a]	6.82 [b]	6.93 [c]	6.81 [d]	7.06 [e]	7.13 [f]	7.31 [g]	7.24 [h]
Salesperson Company Relationship	6.98	6.45 [a]	7.06 [a]	7.12 [b]	6.92 [c]	6.90 [d]	6.77 [e]	7.19 [f]	7.13 [g]
Product Differences	5.18	4.19 [a]	4.61 [b]	5.20 [c]	4.98 [d]	5.27 [e]	5.35 [f]	5.56 [g]	5.56 [h]
Know More	4.99	5.37 [a]	5.45 [b]	4.90 [c]	4.87 [d]	5.18 [e]	5.08 [f]	4.71 [g]	4.87 [h]
Information Differences	6.40	6.15 [a]	6.25 [b]	6.28 [c]	6.23 [d]	6.49 [e]	6.70 [f]	6.56 [g]	6.49 [h]

Table 5-21 Continued

Brand Loyalty Question (yes/no questions where 1 is yes and 0 is no)									
Loyal to Brand	0.48	0.02 [d]	0.05 [e]	0.95 [a]	0.73 [a]	0.10 [b]	0.38 [b]	0.58 [c]	0.45 [c]
Recommend Brand	0.46	0.16 [a]	0.10 [a]	0.99 [b]	0.79 [b]	0.09 [c]	0.32 [c]	0.42 [d]	0.34 [d]
Would Switch Brands for a 5% Discount	0.43	0.87 [a]	0.47 [a]	0.94 [c]	0.92 [d]	0.03 [b]	0.10 [b]	0.07 [e]	0.06 [f]
Would Switch Brands for a 10% Discount	0.63	0.49 [a]	0.75 [a]	1.00 [b]	0.87 [b]	0.74 [c]	0.79 [d]	0.23 [e]	0.24 [f]

Table 5-21 Continued

Company Loyalty Question (yes/no questions where 1 is yes and 0 is no)									
Loyal to Company	0.60	0.05 [d]	0.04 [e]	0.97 [a]	0.82 [a]	0.20 [b]	0.44 [b]	0.87 [c]	0.70 [c]
Recommend Company	0.54	0.16 [a]	0.06 [a]	0.99 [b]	0.84 [b]	0.13 [c]	0.34 [c]	0.66 [d]	0.54 [d]
Would Switch Companies for a 5% Discount	0.42	0.94 [a]	0.54 [a]	0.93 [b]	0.88 [b]	0.02 [c]	0.10 [c]	0.02 [d]	0.04 [e]
Would Switch Companies for a 10% Discount	0.60	0.53 [a]	0.78 [a]	1.00 [b]	0.89 [b]	0.77 [c]	0.92 [c]	0.06 [d]	0.02 [d]

Note: A matching letter indicates that the means for these variables in the same row and same segment are statistically different using T Tests for the mean comparison, Mood's Median Tests for the median comparison, and Two Proportion Z-Tests for proportion comparisons.
** One outlier variable dropped for farm size*

In the Economic buyer segment, the clusters predicted by the challenge model on average placed higher importance on Service Quality, Relationship With Salesperson, and lower importance on Multiple Brands than clusters predicted by the base model in the Dealer/Retailer Attribute Question. In the Information Importance Question clusters predicted by the challenge model on average placed higher importance on information about saving money and lower importance on information about product options than the clusters predicted by the base model. In the Crop Protection Attribute Question clusters predicted by the challenge model on average placed higher importance on price and lower importance on relationship than the clusters predicted by the base model. In the Likert Scale Agreement Question clusters predicted by the challenge model on average had a higher level of agreement with the Salesperson Company Relationship statement than the clusters predicted by the base model. In the Brand Loyalty Question clusters predicted by the challenge model on average had a higher level of agreement to Would Switch Brands for a 10% Discount and a lower level of agreement to Recommend Brand and Would Switch Brands for a 5% Discount than the clusters predicted by the base model. In the Company Loyalty Question clusters predicted by the challenge model on average had a higher level of agreement to Would Switch Companies for a 10% Discount and a lower level of agreement to Recommend Company and Would Switch Companies for a 5% Discount than the clusters predicted by the base model.

In the Agronomic buyer segment, the clusters predicted by the challenge model on average placed higher importance on Credibility and a lower importance on Caring than the clusters predicted by the base model in the Salesperson Characteristics Question. In the Crop Protection Attribute Question clusters predicted by the challenge model on average placed lower importance on relationship than the clusters predicted by the base model. In the Brand Loyalty Question clusters predicted by the challenge model on average had a lower level of agreement to Loyal to Brand, Recommend Brand, and Would Switch Brands for a 10% Discount than the clusters predicted by the base model. In the Company Loyalty Question clusters predicted by the challenge model on average had a lower level of agreement to Recommend Company, Would Switch Companies for a 5% Discount, and Would Switch Companies for a 10% Discount than the clusters predicted by the base model.

In the Business buyer segment, the clusters predicted by the challenge model on average placed higher importance on Availability of Service and lower importance on Multiple Brands

than clusters predicted by the base model in the Dealer/Retailer Attribute Question. In the Product Selection Factor Question clusters predicted by the challenge model on average placed higher importance on Brand and lower importance on Price of Product than the clusters predicted by the base model. In the Salesperson Characteristics Question clusters predicted by the challenge model on average placed higher importance on Relationship and lower importance on Credibility than the clusters predicted by the base model. In the Information Importance Question clusters predicted by the challenge model on average placed a higher importance on information about product options and a lower importance on information about saving money. In the Crop Protection Attribute Question clusters predicted by the challenge model on average placed higher importance on relationship and lower importance on price and performance than the clusters predicted by the base model. In the Brand Loyalty Question clusters predicted by the challenge model on average had a higher level of agreement to Loyal to Brand, Recommend Brand, and Would Switch Brands for a 5% Discount than the clusters predicted by the base model. In the Company Loyalty Question clusters predicted by the challenge model on average had a higher agreement to Recommend Company, Would Switch Companies for a 5% Discount, and Would Switch Companies for a 10% Discount than the clusters predicted by the base model.

In the Relationship buyer segment, the clusters predicted by the challenge model on average placed lower importance on relationship than clusters predicted by the base model in the Crop Protection Attributes Question. In the Brand Loyalty Question clusters predicted by the challenge model on average had a lower level of agreement to Loyal to Brand and Recommend Brand than the clusters predicted by the base model. In the Company Loyalty Question clusters predicted by the challenge model on average had a lower level of agreement to Recommend brand and Would Switch Companies for a 10% Discount than the clusters predicted by the base model.

The number of statistically different means for the segmentation questions by buyer segment is reflective of the challenge model's ability to predict producers into the correct buying segment with the Relationship buyer segment having the lowest number of statistically different means and the highest accurate predictability.

The demographic variables across all buyer segments except the Economic buyer segments are not statistically different. In the Economic buyer segment, the challenge model predicts the proportions for the education categories statistically different than the base model with a higher proportion of respondents with High School or Less, Less than Bachelor's Degree, and Graduate

Degree with a lower proportion of respondents with 4 Year Degree as their highest education. The challenge model also predicts the average age of respondents to be higher than the base model.

5.8 Summary

This chapter presented the results of the phase approach used in this research consisting of cluster analysis and a multinomial logistic regression model for the farmer data in Phase I, cluster analysis and a multinomial logistic regression model for the retailer data in Phase II, identifying farmer segments using the producer multinomial logistic regression model in Phase III, identifying retailers' perceptions of farmer segments using the retailer multinomial logistic regression in phase IV, a multi-level confusion matrix with accuracy measures in Phase V, and a comparison of the results between the farmers and retailers to identify and understand what variables are driving the gaps agricultural retailers' understanding of value proposition for large commercia producers. The final chapter provides interpretation and discussion of results in the light of the original objectives and hypotheses.

CHAPTER 6. CONCLUSION

In this chapter the analyses and results will be summarized to address the motivation and objectives of this research as well as the most important findings that led to the conclusions for the proposed hypotheses. Then, implications around how this research could be used by agricultural retailers will be discussed. Lastly, limitations and further research will be considered.

6.1 Motivation and Objectives

This research was motivated by the increasingly more important need for retailers to maximize their resources and address the producers' buying preferences in order to stay relevant as farms consolidate and become more sophisticated along with a more competitive market. The objective of this study is to identify how demographics and other seller attributes explain differences between retailers' and producers' perception of value created for large commercial producers.

6.2 Proposed Hypotheses

Based on prior research done around agricultural market segmentation and market knowledge, two hypotheses were proposed for this research:

- Hypothesis 1: Significant gaps exist in the buying segments derived from producers expressed buying preferences relative to segments derived from Retailers' perceptions of large producer buying behaviors
- Hypothesis 2: Retailers' perceptions of farmer buying behavior results in significant differences in the sizes of producer buying segments relative to the size of buying segments identified using actual producer data

6.3 Discussion of Key Findings

The key findings of this research were centered around identifying the gaps in farmers buying preferences between segments predicted using farmer buying preferences compared to

segments predicted using retailers' perceived buying preferences and accuracy at which retailers could predict the correct buying segment for their producers.

The comparison of the variables statistically significant between the producer multinomial logistic regression model estimated in Phase I and the retailer multinomial logistic regression model estimated in Phase II for each buyer segment compared to the Relationship buyer segment gave insight into which variables differed in statistical significance or by direction of impact between the two models.

In the models predicting the Economic buyer segment compared to the Relationship buyer segment, the variables Retailer Differences and Product Differences were statistically significant in the producer multinomial logistic regression model but not in the retailer multinomial logistic regression model. The variable Price of Product was statistically significant in the retailer multinomial logistic regression model but not in the producer multinomial logistic regression model.

In the models predicting the Agronomic buyer segment compared to the Relationship buyer segment, the variables Relationship With Salesperson, Past Experience, Salesperson Company Relationship, and Recommend Brand were statistically significant in the producer multinomial logistic regression but not in the retailer multinomial logistic regression model. The variables CP Price, CP Performance, Loyal to Company, Reliability, Credibility, Information Save Money, and Effectiveness were statistically significant in the retailer multinomial logistic regression model, but not in the producer multinomial logistic regression model.

In the models predicting the Business buyer segment compared to the Relationship buyer segment, the variables Service Quality, Relationship With Salesperson, Multiple Brands, CP Price, Loyal to Brand, Recommend Brand, Would Switch Brands for a 5% Discount, Would Switch Brands for a 10% Discount, Loyal to Company, Reliability, and Information Product Options were statistically significant in the producer multinomial logistic regression model but not in the retailer multinomial logistic regression model.

The variables that differed in statistical significance or direction of impact between the producer multinomial logistic regression model and the retailer multinomial logistic regression model gave further explanation into which variables made up the gap between the farmer and retailer buying segments. This gap did not seem to have large impact on the proportions of each segment determined by farmer data with the producer multinomial logistic regression estimates

(farmer predicted farmer clusters) and by farmer data with the retailer multinomial logistic regression estimates (retailer predicted farmer clusters) when compared in section 5.7.2. The segments in order of size were consistent between the farmer predicted farmer clusters and the retailer predicted retailer clusters. However, the accuracy measures from the multilevel confusion matrix in Phase V concluded that these variables causing the gaps between the producer multinomial logistic regression model and the retailer multinomial logistic regression model did have a negative impact on retailers' ability to predict their producers in the correct buying segment. These results served as the evidence needed to address the proposed hypotheses.

6.4 Summary of Hypothesis Test

Based on the analysis we cannot reject Hypothesis 1 and cannot reject Hypothesis 2. Hypothesis 1 cannot be rejected due to differences in statistically significant variables derived from producers expressed buying preferences relative to retailers perceived buying preferences in the multinomial logistic regression models as well as differences in accuracy measures between segments. Hypothesis 2 cannot be rejected due to buyer segment proportions determined by the producer multinomial logistic regression model using farmer data and retailer multinomial logistic regression model using farmer data being statistically different for the Agronomic and Business buyer segments.

6.5 Implications

This research could be the foundation for directing marking efforts for agricultural retailers by choosing the correct messages to appeal to each buying segment. Increased customer satisfaction could lead to elevated sales performance and higher customer retention. As retailers continue to consolidate, growing both their geography and number of customers served, the need for more sophisticated marketing techniques such as the cluster analysis performed here will help refine the value proposition to farmer customers. Salespeople were on average quite accurate in their assessment of different farmer segments in this study. The results here suggest that a retailer would do well to utilize their sale people as a key resource in identifying the right farmer segments to serve.

6.6 Limitations of the Study and Further Research

The limitations of this study primarily occurred due to the data available. There was not a way to connect the retailers to the producers that shaped their perception of added value and importance around farmers buying behaviors. While the ranking question format offers a significant amount of information, it limited the insights that could be gathered from the multinomial logistic regression model due to the inability for ceteris paribus preventing the use of marginal effects.

Further research can be explored about the differences in these segments for other services such as seed, fertilizer, and animal products to offer a more diverse segmentation of producers as well as differences in accuracy measures between retailer respondents at varying levels in the company.

REFERENCES

Aldenderfer, M. S., & Blashfield, R. K. (2011). A Review of Clustering Methods. In M. S. Aldenderfer, & R. K. Blashfield, *Cluster Analysis* (pp. 34-62). Thousand Oaks: SAGE Publications, Inc.

Alexander, C. E., Wilson, C. A., & Foley, D. H. (2005). Agricultural Input Market Segments: Who Is Buying What? *Journal of Agribusiness, 23, 2*, 113-132.

Alexander, C., Boehlje, M., Downey, W. S., Gray, A., Gunderson, M., & Roucan-Kane, M. (2009). *Serving Producers in Volatile Times: Themes from the 2008 Large Commercial Producer Survey.*

Almaliki, Z. A. (2018, July 27). *Standardization VS Normalization*. Retrieved from Medium: https://dataakkadian.medium.com/standardization-vs-normalization-da7a3a308c64

Baker, R. J. (2006). *Pricing on Purpose: Creating and Capturing Value.* Hoboken: John Wiley & Sons, Inc.

Borchers, B., Roucan-Kane, M., Alexander, C., Boehlje, M. D., Downey, W. S., & Gray, A. W. (2012). How Large Commercial Producers Choose Input Suppliers: Expendable Products from Seed to Animal Health. *International Food and Agribusiness Management Review, Volume 15, Issue 2*.

Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2012). *NbClust Package. An examination of indices for determining the number of clusters.* Retrieved from HAL open science: https://hal.archives-ouvertes.fr/hal-01126138/document#:~:text=NbClust%20package%20provides%2030%20indices,distance%20measures%2C%20and%20clustering%20methods.&text=diss%20dissimilarity%20matrix%20to%20be%20used.

Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software, Volume 61, Issue 6*.

Coefficient of Variation: Definition, Formula, Interpretation, Examples & FAQs. (n.d.). Retrieved from formplus: https://www.formpl.us/blog/coefficient-variation

Donnelly, R. A., & Abdel-Raouf, F. (2016). *Statistics, 3E.* Indianapolis: DK Publishing.

Eggert, A., Ulaga, W., Frow, P., & Payne, A. (2018). Conceptualizing and communicating value in business markets: From value in exchange to value in use. *Industrial Marketing Management, 69*, 80-90.

Farming and Farm Income. (2022, February 4). Retrieved from United States of Agriculture, Economic Research Service: https://www.ers.usda.gov/data-products/ag-and-food-statistics-charting-the-essentials/farming-and-farm-income/

Gloy, B. A., & Akridge, J. T. (1999). Segmenting the Commercial Producer Marketplace for Agricultural Inputs. *International Food and Agribusiness Management Review, 2(2)*, 145-163.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis.* Hampshire: Cengage Learning.

Harbor, A. L., Martin, M. A., & Akridge, J. T. (2008). Assessing Input Brand Loyalty among U.S. Agricultural Producers. *International Food and Agribusiness Management Review, Volume 11, Issue 1*.

Hoppe, R. (2015, February 2). *Profit Margin Increases With Farm Size*. Retrieved from United States Department of Agriculture, Economic Research Service: https://www.ers.usda.gov/amber-waves/2015/januaryfebruary/profit-margin-increases-with-farm-size/

Hoppe, R. (2017, December 30). *Profit margins tend to fluctuate with farm size*. Retrieved from Iowa Farmer Today: https://www.agupdate.com/iowafarmertoday/opinion/profit-margins-tend-to-fluctuate-with-farm-size/article_2fe5dd54-ea8a-11e7-b41e-cf0112044cc8.html#:~:text=Farming%20is%20still%20overwhelmingly%20comprised,operate%20half%20of%20the%20farmland.

Hoppe, R. A., & MacDonald, J. M. (2013, April). *Updating the ERS Farm Typology.* Retrieved from U.S. Department of Agriculture, Economic Research Service: https://www.ers.usda.gov/webdocs/publications/43742/36482_eib110.pdf?v=5529.2

Ketchen, D. J., & Shook, C. L. (1996). The Application of Cluster Analysis in Strategic management Research: An Analysis and Critique. *Strategic management Journal, Volume 17*, 441-458.

Lai, J., Olynk Widmar, N. J., Gunderson, M. A., Widmar, D. A., & Ortega, D. L. (2018). Prioritization of farm success factors by commercial farm managers. *International Food and Agribusiness Management Review, Volume 21, Issue 6*.

MacDonald, J. M. (2020). Tracking the Consolidation of U.S. Agriculture. *Applied Economic Perspectives and Policy, Volume 42, Number 3*, 361-379.

Milligan, G. W. (1980). An Examination of the Effects of Six Types of Error Perturbation on Fiften Clustering Algorithms. *Psychometrika, Volume 45, Number 3*.

mlogit - Multinomial (polytomous) logistic regression. (n.d.). Retrieved from Stata.com: https://www.stata.com/manuals13/rmlogit.pdf

Mohajon, J. (2020, May 29). *Confusion Matrix for Your Multi-Class Machine Learning Model*. Retrieved from Towards Data Science: https://towardsdatascience.com/confusion-matrix-for-your-multi-class-machine-learning-model-ff9aa3bf7826

Payne, A., Frow, P., & Eggert, A. (2017). The customer value proposition: evoluation, development, and application in marketing. *Journal of the Academy of Marketing Science, 45*, 467-489.

Ratick, S. J., & Schwarz, G. (2009). Monte Carlo Simulation. In N. Castree, M. Crang, & M. Domosh, *International Encyclopedia of Human Geography* (pp. 175-184). Amsterdam: Elsevier.

Ravald, A., & Grönroos, C. (1996). The value concept and relationship marketing. *European Journal of Marketing, Volume 30, Number 2*, 19-30.

Reimer, A., Downey, W. S., & Akridge, J. (2009). Market Segmentation Practices of Retail Crop Input Firms. *International Food and Agribusiness Management Review, Volume 12, Issue 1*.

Roucan-Kane, M., Alexander, C., Boehlje, M. D., Downey, S. W., & Gray, A. W. (2011). Large Commercial Producer Market Segments for Agricultural Capital Equipment. *International Food and Agribusiness Management Review, Volume 14, Issue 4*.

Terho, H., Eggert, A., Haas, A., & Ulaga, W. (2015). How sales strategy translates into performance: The role of salesperson customer orientation and value-based selling. *Industrial Marketing Management, 45*, 12-21.

The CLUSTER Procedure. (2017). In *SAS/STAT 14.3 User's Guide* (pp. 2194-2277). Cary: SAS Institute Inc.

Timberlake, J. L. (2012). Structural change in agriculture: Analyzing strategic accounts and retailer relationships. West Lafayette, Indiana, United States.

Ulaga, W., & Eggert, A. (2006). Value-Based Differentiation in Business Relationships: Gaining and Sustaining Key Supplier Status. *Journal of Marketing, Volume 70*, 119-136.

APPENDIX A. SURVEY QUESTIONS

Farmer Survey

Please rank the **top three most important** factors when selecting which dealer/retailer you purchase products from. *(1 is the top choice, then 2 and 3)*

_____ Availability of Multiple Product Brands

_____ Availability of Service (application, delivery, etc.)

_____ Best Value or Price

_____ Location of the Retailer Facilities

_____ Past Experience with the Supplier

_____ Product Returns and Warranties

_____ Quality of Service Performed

_____ Relationship with Salesperson there

In general, which of the following factors is most important in your selection of what products to use? Please rank from 1 to 6, where 1 is *"the most important"* and 6 is *"the least important"*. *(Drag each term and drop it to desired rank)*

_____ Availability of product

_____ Brand of product

_____ Ease of handling the product

_____ Price of product

_____ Effectiveness of product

_____ Other non-price qualities of product

In general, how would you rank the following characteristics of the salespeople you choose to work with? Please rank from 1 to 4, where 1 is *"the most important"* and 4 is *"the least important"*. *(Drag each term and drop it to desired rank)*

_____ Relationship - you can trust them with confidential information

_____ Reliability - you can trust the salesperson with follow through

_____ Credibility - you can trust the information the salesperson provides

_____ Caring - they care about my success as much as their own

Thinking about the last time you bought inputs (seed, fertilizer, chemical, feed, animal health, and etc.) from a dealer/retailer, which of the following types of information was most important to you during that purchase? Please rank from 1 to 3, where 1 is *"the most important"* and 3 is *"the least important"*. *(Drag each term and drop it to desired rank)*

_____ Options for how to save money

_____ Information about different product options

_____ Advice about how to maximize profits

For each of the purchases below, please rank these attributes in order of importance to you, where 1 is ***"most important"*** and 3 is ***"least important"***. *(put a 1, 2, and 3 in each column)*

	Price	Product Performance	Supplier Relationship
Seed			
Fertilizer			
Crop Protection			
Capital Equipment			

Using a scale of 1 to 9, where 1 means ***"strongly disagree"*** and 9 means ***"strongly agree"***, how do you feel about the following? *(use the slider to choose your scale)*

	1 2 3 4 5 6 7 8 9
Significant differences exist in the quality of services between different dealers and retailers	
My relationship with sales people is more important than the relationship I have with the company they represent	
Significant differences exist between generic expendable products *(seed, feed/nutrition, etc.)* and branded products	
I know more about my expendable	

products than my dealer or retailer	
Significant differences exist in the quality of information I receive from different dealers and retailers	

For the **BRANDS** from which you ***primarily*** purchase the products listed below, please indicate if you agree with the following statements: *(Check all that apply)*

	Seed	**Fertilizer**	**Crop Protection**	**Capital Equipment**
I consider myself loyal to my primary brand in these product categories	☐	☐	☐	☐
I have recommended this brand to someone else	☐	☐	☐	☐
I would switch from this brand for a 5% discount in these product categories	☐	☐	☐	☐
I would switch from this brand for a 10% discount in these product categories	☐	☐	☐	☐

For the **DEALER/RETAILER** from which you ***primarily*** purchase the products listed below, please indicate if you agree with the following statements: *(Check all that apply)*

	Seed	**Fertilizer**	**Crop Protection**	**Capital Equipment**
I consider myself loyal to the primary dealer/retailer I purchase these product categories	☐	☐	☐	☐
I have recommended this dealer/retailer to someone else	☐	☐	☐	☐
I would switch from this dealer/retailer for a 5% discount on these product categories	☐	☐	☐	☐
I would switch from this dealer/retailer for a 10% discount on these product categories	☐	☐	☐	☐

Retailer Survey

Rank the following factors in the order of most importance to large producers in your area when they select the dealer/retailer they purchase products from? Drag and drop the options in the list below so that ***"the most important"*** is on top ***"the least important"*** is on the bottom.

_____ Quality of service performed
_____ Availability of service (application, delivery, etc.)
_____ Relationship with salesperson
_____ Availability of multiple brands
_____ Past experience with your business
_____ Product returns and warranties

In general, which of the following factors do you think is most important to large producers in your area in selecting which products to use? Drag and drop the options in the list below so that ***"the most important"*** is on top ***"the least important"*** is on the bottom.

_____ Availability of product
_____ Brand of product
_____ Ease of handling the product
_____ Price of product
_____ Effectiveness of product
_____ Other non-price qualities of product

In general, what is your perception of how large producers in your area rank the following characteristics of the salespeople they choose to work with? Drag and drop the options in the list below so that ***"the most important"*** is on top ***"the least important"*** is on the bottom.

_____ Relationship - they can trust the salesperson with confidential information
_____ Reliability - they can trust the salesperson will follow through
_____ Credibility - they can trust the information the salesperson provides
_____ Caring - the salesperson cares as much about the customers success as their own

Thinking about the last time your large farmer customers bought inputs from you, rank the types of information you think were most important to him/her during that purchase? Drag and drop the options in the list below so that ***"the most important"*** is on top ***"the least important"*** is on the bottom.

_____ Options for how to save money
_____ Information about different product options
_____ Advice about how to maximize profits

For each of the purchases listed below, please rank these attributes: **"Price"**, **"Product Performance"**, and **"Dealer/Retailer Relationship"** in order of importance to large producers in your area, with 1 being the ***"most important attribute"*** and 3 being ***"the least important attribute"***.

	Price	Product Performance	Dealer/Retailer Relationship
Seed			
Crop Protection			
Fertilizer			

Using a scale of ***"Very Strongly Disagree"*** to ***"Very Strongly Agree"***, indicate your level of agreement with the following statements:

	Very Strongly Disagree	2	3	4	Neither Agree nor Disagree	6	7	8	Very Strongly Agree
Large producers believe that significant differences exist in the quality of services between different dealers/retailers									
Large producers believe that their relationship with the sales people is more important than the relationship they have with the company they represent									
Large producers believe that significant differences exist between generic expendable products (seed, feed/nutrition, etc.) and branded products									

Large producers believe that they know more about the products they purchase than the dealer/retailer	
Large producers believe that significant differences exist in the quality of information they receive from different dealers/retailers	

Thinking about the product **BRANDS** that your large farmer customers ***primarily purchase***, please indicate if you agree with the following statements for each product: *(check all that apply)*

	Seed	Fertilizer	Crop Protection
I believe my large customers are loyal to the brands they purchase	•	•	•
I believe my large customers recommend the brands they purchase to others	•	•	•
I believe my large customers would switch brands for a 5% discount	•	•	•
I believe my large customers would switch brands for a 10% discount	•	•	•

Thinking about the **COMPANY** you represent, please indicate if you **agree** with the following statements about your large farmer customers' in relation to each product. *(check all that apply)*

	Seed	Fertilizer	Crop Protection
I believe my large customers are loyal to my company	•	•	•
I believe my large customers recommend my company to others	•	•	•
I believe my large customers would switch from my company for a 5% discount	•	•	•
I believe my large customers would switch from my company for a 10% discount	•	•	•

APPENDIX B. TABLES

Table B-1 Service Quality (Farmer Survey)

	Freq.	Percent
1	187	14.42
2	217	16.73
3	226	17.42
4	667	51.43
Total	1297	100

Table B-2 Availbility of Service (Farmer Survey)

	Freq.	Percent
1	111	8.56
2	182	14.03
3	170	13.11
4	834	64.30
Total	1,297	100

Table B-3 Relationship With Salesperson (Farmer Survey)

	Freq.	Percent
1	159	12.26
2	174	13.42
3	174	13.42
4	790	60.91
Total	1,297	100

Table B-4 Multiple Brands (Farmer Survey)

	Freq.	Percent
1	52	4.01
2	73	5.63
3	89	6.86
4	1,083	83.50
Total	1,297	100

Table B-5 Past Experience (Farmer Survey)

	Freq.	Percent
1	149	11.49
2	222	17.12
3	219	16.89
4	707	54.51
Total	1,297	100

Table B-6 Product Returns and Warranties (Farmer Survey)

	Freq.	Percent
1	17	1.31
2	56	4.32
3	81	6.25
4	1,143	88.13
Total	1,297	100

Table B-7 Availability (Farmer Survey)

	Freq.	Percent
1	112	8.64
2	194	14.96
3	435	33.54
4	354	27.29
5	161	12.41
6	41	3.16
Total	1,302	100

Table B-8 Brand (Farmer Survey)

	Freq.	Percent
1	41	3.16
2	66	5.09
3	106	8.17
4	211	16.27
5	465	35.85
6	408	31.46
Total	1,297	100

Table B-9 Ease of Handling (Farmer Survey)

	Freq.	Percent
1	16	1.23
2	114	8.79
3	334	25.75
4	440	33.92
5	315	24.29
6	78	6.01
Total	1,297	100

Table B-10 Price of Product (Farmer Survey)

	Freq.	Percent
1	283	21.82
2	628	48.42
3	236	18.20
4	102	7.86
5	31	2.39
6	17	1.31
Total	1,297	100

Table B-11 Effectiveness (Farmer Survey)

	Freq.	Percent
1	837	64.53
2	254	19.58
3	101	7.79
4	61	4.70
5	33	2.54
6	11	0.85
Total	1,297	100

Table B-12 Non-Price Qualities (Farmer Survey)

	Freq.	Percent
1	8	0.62
2	41	3.16
3	85	6.55
4	129	9.95
5	292	22.51
6	742	57.21
Total	1,297	100

Table B-13 Relationship (Farmer Survey)

	Freq.	Percent
1	293	22.59
2	216	16.65
3	361	27.83
4	427	32.92
Total	1,297	100

Table B-14 Reliability (Farmer Survey)

	Freq.	Percent
1	338	26.06
2	475	36.62
3	354	27.29
4	130	10.02
Total	1,297	100

Table B-15 Credibility (Farmer Survey)

	Freq.	Percent
1	491	37.86
2	386	29.76
3	282	21.74
4	138	10.64
Total	1,297	100

Table B-16 Caring (Farmer Survey)

	Freq.	Percent
1	175	13.49
2	220	16.96
3	300	23.13
4	602	46.41
Total	1,302	100

Table B-17 Information Save Money (Farmer Survey)

	Freq.	Percent
1	397	30.61
2	519	40.02
3	381	29.38
Total	1,297	100

Table B-18 Information Product Options (Farmer Survey)

	Freq.	Percent
1	375	28.91
2	320	24.67
3	602	46.41
Total	1,297	100

Table B-19 Information Maximize Profits (Farmer Survey)

	Freq.	Percent
1	525	40.48
2	458	35.31
3	314	24.21
Total	1,297	100

Table B-20 CP Price (Farmer Survey)

	Freq.	Percent
1	382	29.45
2	646	49.81
3	269	20.74
Total	1,297	100

Table B-21 CP Performance (Farmer Survey)

	Freq.	Percent
1	747	57.59
2	394	30.38
3	156	12.03
Total	1,297	100

Table B-22 CP Relationship (Farmer Survey)

	Freq.	Percent
1	168	12.95
2	257	19.81
3	872	67.23
Total	1,297	100

Table B-23 Retailer Differences (Farmer Survey)

	Freq.	Percent
1	26	2.00
2	20	1.54
3	34	2.62
4	41	3.16
5	160	12.34
6	153	11.80
7	263	20.28
8	267	20.59
9	333	25.67
Total	1,297	100

Table B-24 Salesperson Company Relationship (Farmer Survey)

	Freq.	Percent
1	17	1.31
2	29	2.24
3	37	2.85
4	38	2.93
5	134	10.33
6	147	11.33
7	249	19.20
8	332	25.60
9	314	24.12
Total	1,297	100

Table B-25 Product Differences (Farmer Survey)

	Freq.	Percent
1	53	4.09
2	91	7.02
3	124	9.56
4	132	10.18
5	294	22.67
6	191	14.73
7	195	15.03
8	133	10.25
9	84	6.48
Total	1,297	100

Table B-26 Know More (Farmer Survey)

	Freq.	Percent
1	55	4.24
2	93	7.17
3	159	12.26
4	189	14.57
5	387	29.84
6	182	14.03
7	104	8.02
8	86	6.63
9	42	3.24
Total	1,297	100

Table B-27 Information Differences (Farmer Survey)

	Freq.	Percent
1	14	1.08
2	36	2.78
3	50	3.86
4	56	4.32
5	276	21.28
6	217	16.73
7	232	17.89
8	218	16.81
9	198	15.27
Total	1,297	100

Table B-28 Loyal to Brand (Farmer Survey)

	Freq.	Percent
0	662	51.04
1	635	48.96
Total	1,297	100

Table B-29 Recommend Brand (Farmer Survey)

	Freq.	Percent
0	695	53.59
1	602	46.41
Total	1,297	100

Table B-30 Would Switch Brands for a 5% Discount (Farmer Survey)

	Freq.	Percent
0	769	59.29
1	528	40.71
Total	1,297	100

Table B-31 Would Switch Brands for a 10% Discount (Farmer Survey)

	Freq.	Percent
0	488	37.63
1	809	62.37
Total	1,297	100

Table B-32 Loyal to Company (Farmer Survey)

	Freq.	Percent
0	518	39.94
1	779	60.02
Total	1,297	100

Table B-33 Recommend Company (Farmer Survey)

	Freq.	Percent
0	622	47.96
1	675	52.04
Total	1,297	100

Table B-34 Would Switch Companies for a 5% Discount (Farmer Survey)

	Freq.	Percent
0	769	59.29
1	528	40.71
Total	1,297	100

Table B-35 Would Switch Companies for a 10% Discount (Farmer Survey)

	Freq.	Percent
0	522	40.25
1	775	59.75
Total	1,297	100

Table B-36 Service Quality (Retailer Survey)

	Freq.	Percent
1	134	24.14
2	211	38.02
3	138	24.86
4	72	12.97
Total	555	100

Table B-37 Availability of Service (Retailer Survey)

	Freq.	Percent
1	295	53.15
2	119	21.44
3	85	15.32
4	56	10.09
Total	555	100

Table B-38 Relationship With Salesperson (Retailer Survey)

	Freq.	Percent
1	80	14.41
2	111	20
3	161	29.01
4	203	36.58
Total	555	100

Table B-39 Multiple Brands (Retailer Survey)

	Freq.	Percent
1	6	1.08
2	17	3.06
3	29	5.23
4	503	90.63
Total	555	100

Table B-40 Past Experience (Retailer Survey)

	Freq.	Percent
1	37	6.67
2	96	17.30
3	121	21.80
4	301	54.23
Total	555	100

Table B-41 Product Returns and Warranties (Retailer Survey)

	Freq.	Percent
1	3	0.54
2	1	0.18
3	21	3.78
4	530	95.5
Total	555	100

Table B-42 Availability (Retailer Survey)

	Freq.	Percent
1	21	3.78
2	82	14.77
3	155	27.93
4	132	23.78
5	118	21.26
6	47	8.47
Total	555	100

Table B-43 Brand (Retailer Survey)

	Freq.	Percent
1	18	3.24
2	26	4.68
3	61	10.99
4	91	16.4
5	173	31.17
6	186	33.51
Total	555	100

Table B-44 Ease Handling (Retailer Survey)

	Freq.	Percent
1	20	3.6
2	71	12.79
3	150	27.03
4	179	32.25
5	91	16.4
6	44	7.93
Total	555	100

Table B-45 Price of Product (Retailer Survey)

	Freq.	Percent
1	170	30.63
2	210	37.84
3	91	16.4
4	55	9.91
5	18	3.24
6	11	1.98
Total	555	100

Table B-46 Effectiveness (Retailer Survey)

	Freq.	Percent
1	324	58.38
2	144	25.95
3	59	10.63
4	18	3.24
5	6	1.08
6	4	0.72
Total	555	100

Table B-47 Non-Price Qualities (Retailer Survey)

	Freq.	Percent
1	2	0.36
2	22	3.96
3	39	7.03
4	80	14.41
5	149	26.85
6	263	47.39
Total	555	100

Table B-48 Relationship (Retailer Survey)

	Freq.	Percent
1	167	30.09
2	104	18.74
3	139	25.05
4	145	26.13
Total	555	100

Table B-49 Reliability (Retailer Survey)

	Freq.	Percent
1	153	27.57
2	190	34.23
3	140	25.23
4	72	12.97
Total	555	100

Table B-50 Credibility (Retailer Survey)

	Freq.	Percent
1	169	30.45
2	180	32.43
3	124	22.34
4	82	14.77
Total	555	100

Table B-51 Caring (Retailer Survey)

	Freq.	Percent
1	66	11.89
2	81	14.59
3	152	27.39
4	256	46.13
Total	555	100

Table B-52 Information Save Money (Retailer Survey)

	Freq.	Percent
1	132	23.78
2	190	34.23
3	233	41.98
Total	555	100

Table B-53 Information Product Options (Retailer Survey)

	Freq.	Percent
1	139	25.05
2	190	34.23
3	226	40.72
Total	555	100

Table B-54 Information Maximize Profits (Retailer Survey)

	Freq.	Percent
1	284	51.17
2	175	31.53
3	96	17.3
Total	555	100

Table B-55 CP Price (Retailer Survey)

	Freq.	Percent
1	199	35.86
2	181	32.61
3	175	31.53
Total	555	100

Table B-56 CP Performance (Retailer Survey)

	Freq.	Percent
1	249	44.86
2	217	39.1
3	89	16.04
Total	555	100

Table B-57 CP Relationship (Retailer Survey)

	Freq.	Percent
1	107	19.28
2	157	28.29
3	291	52.43
Total	555	100

Table B-58 Retailer Differences (Retailer Survey)

	Freq.	Percent
1	12	2.16
2	27	4.86
3	35	6.31
4	30	5.41
5	60	10.81
6	89	16.04
7	144	25.95
8	82	14.77
9	76	13.69
Total	555	100

Table B-59 Salesperson Company Relationship (Retailer Survey)

	Freq.	Percent
1	19	3.42
2	16	2.88
3	15	2.7
4	17	3.06
5	50	9.01
6	67	12.07
7	112	20.18
8	114	20.54
9	145	26.13
Total	555	100

Table B-60 Product Differences (Retailer Survey)

	Freq.	Percent
1	32	5.77
2	41	7.39
3	74	13.33
4	52	9.37
5	125	22.52
6	85	15.32
7	81	14.59
8	41	7.39
9	24	4.32
Total	555	100

Table B-61 Know More (Retailer Survey)

	Freq.	Percent
1	36	6.49
2	73	13.15
3	101	18.2
4	85	15.32
5	113	20.36
6	61	10.99
7	46	8.29
8	21	3.78
9	19	3.42
Total	555	100

Table B-62 Information Differences (Retailer Survey)

	Freq.	Percent
1	8	1.44
2	20	3.6
3	25	4.5
4	27	4.86
5	83	14.95
6	99	17.84
7	131	23.6
8	105	18.92
9	57	10.27
Total	555	100

Table B-63 Loyal to Brand (Retailer Survey)

	Freq.	Percent
0	323	58.2
1	232	41.8
Total	555	100

Table B-64 Recommend Brand (Retailer Survey)

	Freq.	Percent
0	211	38.02
1	344	61.98
Total	555	100

Table B-65 Would Switch Brands for a 5% Discount (Retailer Survey)

	Freq.	Percent
0	212	38.2
1	343	61.8
Total	555	100

Table B-66 Would Switch Brands for a 10% Discount (Retailer Survey)

	Freq.	Percent
0	71	12.79
1	484	87.21
Total	555	100

Table B-67 Loyal to Company (Retailer Survey)

	Freq.	Percent
0	240	43.24
1	315	56.76
Total	555	100

Table B-68 Recommend Company (Retailer Survey)

	Freq.	Percent
0	201	36.22
1	354	63.78
Total	555	100

Table B-69 Would Switch Companies for a 5% Discount (Retailer Survey)

	Freq.	Percent
0	310	55.86
1	245	44.14
Total	555	100

Table B-70 Would Switch Companies for a 10% Discount (Retailer Survey)

	Freq.	Percent
0	127	22.88
1	428	77.12
Total	555	100

Table B-71 Farmer Clusters

	Freq.	Percent
1	177	13.65
2	370	28.53
3	370	28.53
4	380	29.30
Total	1,297	100

Table B-72 Retailer Clusters

	Freq.	Percent
1	179	32.25
2	114	20.54
3	151	27.21
4	111	20.00
Total	555	100

Table B-73 Farmer Multinomial Logistic Regression Estimates

VARIABLES	(1) 1	(2) 2	(3) 3	(4) 4
Service Quality	11.58	6.530	6.622*	-
	(7.768)	(9.167)	(3.887)	
Availability of Service	-7.150	5.552	5.373	-
	(8.207)	(10.23)	(4.018)	
Relationship With Salesperson	-0.320	33.60***	7.664*	-
	(8.148)	(10.51)	(3.961)	
Multiple Brands	-3.307	17.47	-10.69*	-
	(9.275)	(13.39)	(5.617)	
Past Experience	2.670	20.33**	2.840	-
	(7.132)	(10.05)	(3.767)	
Product Returns and Warranties	-2.437	-17.02	-12.69	-
	(14.54)	(21.80)	(10.32)	
CP Price	-51.72***	-9.074	-47.58***	-
	(13.15)	(21.04)	(7.751)	
CP Performance	-44.47***	-18.79	-34.96***	-
	(12.09)	(18.85)	(6.746)	
CP Relationship	-	-	-	-

Table B-73 Continued

Retailer Differences	-9.580**	-5.108	-0.983	-
	(4.148)	(6.547)	(2.477)	
Salesperson Company Relationship	-2.142	18.01**	-2.061	-
	(3.776)	(7.694)	(2.559)	
Product Differences	-10.32**	-6.883	-0.221	-
	(4.010)	(5.334)	(2.279)	
Know More	1.071	-3.240	0.558	-
	(3.900)	(5.999)	(2.384)	
Information Differences	5.641	-1.817	-2.080	-
	(4.778)	(6.205)	(2.328)	
Loyal to Brand	-179.2***	29.48	-125.9***	-
	(25.56)	(25.50)	(13.12)	
Recommend Brand	-38.03*	93.11***	-54.75***	-
	(19.81)	(27.27)	(12.21)	
Would Switch Brands for a 5% Discount	150.7***	155.7***	45.00*	-
	(24.48)	(33.21)	(26.31)	
Would Switch Brands for a 10% Discount	87.05***	158.8***	125.0***	-
	(20.95)	(28.19)	(11.56)	
Loyal to Company	-200.9***	-19.64	-166.7***	-
	(21.28)	(29.96)	(14.04)	
Recommend Company	-123.8***	49.08	-125.7***	-
	(21.71)	(30.16)	(12.07)	
Would Switch Companies for a 5% Discount	189.6***	184.8***	39.29	-
	(28.57)	(33.81)	(27.61)	
Would Switch Companies for a 10% Discount	130.8***	157.2***	168.4***	-
	(21.29)	(27.75)	(15.27)	
Relationship	3.137	-7.113	5.003	-
	(7.494)	(11.36)	(4.029)	
Reliability	-8.947	-6.300	-8.841*	-
	(9.535)	(14.00)	(5.231)	
Credibility	3.650	-9.596	-1.545	-
	(8.061)	(11.86)	(4.075)	
Caring	-	-	-	-
Information Save Money	-33.18***	11.98	-11.41	-
	(11.92)	(15.43)	(7.305)	
Information Product Options	-3.776	0.258	12.10**	-
	(10.08)	(13.74)	(5.756)	
Information Maximize Profits	-	-	-	-
Availability	5.118	14.79	5.447	-
	(8.214)	(11.87)	(4.868)	

Table B-73 Continued

Brand	8.038 (8.249)	-1.918 (9.733)	7.833 (4.840)	-
Ease of Handling	10.56 (8.955)	-2.557 (12.21)	5.663 (5.170)	-
Price of Product	-5.510 (9.753)	16.90 (13.76)	-1.836 (6.076)	-
Effectiveness	7.025 (9.015)	-23.42 (16.51)	1.578 (5.650)	-
Non-Price Qualities	-	-	-	-
Constant	233.4 (194.7)	-1,430*** (296.9)	291.4** (118.1)	-
Observations	117	370	370	380

Standard errors in parentheses
**** p<0.01, ** p<0.05, * p<0.1*

Table B-74 Retailer Multinomial Logistic Regression Estimates

VARIABLES	(1) 1	(2) 2	(3) 3	(4) 4
Service Quality	-61.30 (42.09)	-0.801 (35.60)	9.784 (30.32)	-
Availability of Service	-40.25 (39.70)	11.97 (34.87)	0.543 (30.08)	-
Relationship With Salesperson	-54.03 (42.35)	-2.317 (36.09)	23.25 (30.15)	-
Multiple Brands	-29.93 (46.32)	21.80 (40.77)	36.54 (36.58)	-
Past Experience	-29.44 (40.29)	9.062 (34.87)	17.25 (29.56)	-
Product Returns and Warranties	-	-	-	-
CP Price	-133.4*** (22.17)	-87.46*** (18.52)	15.81 (15.62)	-
CP Performance	-58.70*** (20.63)	-62.95*** (18.58)	37.14** (16.86)	-
CP Relationship	-	-	-	-

Table B-74 Continued

Retailer Differences	-1.659	-2.580	-0.299	-
	(5.235)	(4.175)	(4.234)	
Salesperson Company Relationship	-1.111	-2.313	-4.027	-
	(5.436)	(4.468)	(4.070)	
Product Differences	-4.936	-0.420	1.393	-
	(4.920)	(4.312)	(4.431)	
Know More	1.340	-0.0960	1.984	-
	(4.485)	(4.114)	(3.590)	
Information Differences	0.777	4.125	3.039	-
	(6.316)	(5.722)	(5.512)	
Loyal to Brand	-62.17***	-19.88	8.165	-
	(23.29)	(17.06)	(15.32)	
Recommend Brand	-58.71***	7.875	18.79	-
	(20.64)	(16.38)	(15.81)	
Would Switch Brands for a 5% Discount	128.8***	148.1***	-1.793	-
	(29.66)	(31.36)	(24.41)	
Would Switch Brands for a 10% Discount	82.61**	40.94*	65.24***	-
	(37.01)	(24.30)	(22.74)	
Loyal to Company	-146.8***	65.10**	28.21	-
	(34.10)	(25.56)	(19.83)	
Recommend Company	-144.6***	-26.26	-58.47**	-
	(27.99)	(21.03)	(23.39)	
Would Switch Companies for a 5% Discount	131.5***	98.72***	3.114	-
	(31.67)	(29.91)	(27.70)	
Would Switch Companies for a 10% Discount	175.4***	160.0***	152.4***	-
	(39.64)	(33.78)	(54.20)	
Relationship	13.37	-5.154	-9.097	-
	(11.45)	(9.086)	(8.481)	
Reliability	3.413	-19.20**	-2.588	-
	(10.35)	(8.836)	(8.471)	
Credibility	17.30	-20.73**	-4.814	-
	(11.59)	(9.453)	(8.398)	
Caring	-	-	-	-
Information Save Money	-63.18***	-21.98*	10.79	-
	(15.65)	(12.76)	(13.21)	
Information Product Options	19.98	2.674	8.853	-
	(12.85)	(10.43)	(11.65)	

Table B-74 Continued

Information Maximize Profits	-	-	-	-
Availability	3.279	-7.047	-7.711	-
	(9.804)	(7.088)	(7.591)	
Brand	-0.904	-5.535	-12.13	-
	(10.53)	(8.240)	(7.936)	
Ease of Handling	4.365	1.605	-5.199	-
	(10.56)	(8.465)	(8.257)	
Price of Product	-21.26*	-10.71	-8.868	-
	(11.58)	(8.460)	(8.988)	
Effectiveness	-8.235	-25.33**	-11.37	-
	(14.22)	(12.17)	(11.34)	
Non-Price Qualities	-	-	-	-
Constant	792.1	-314.3	-598.1	-
	(643.6)	(536.1)	(476.8)	
Observations	179	151	114	111

Standard errors in parentheses
**** p<0.01, ** p<0.05, * p<0.1*