AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Journal of Agricultural and Resource Economics 38(3):500–519 ISSN: 1068-5502 (Print); 2327-8285 (Online)
Copyright 2023 the authors
doi: 10.22004/ag.econ.322852

Selectivity Bias and Cattle Price in the Cattle Procurement Market

Kayla Hildebrand and Chanjin Chung

We examine selectivity bias in the US cattle procurement market. We hypothesize that feedlots optimize profits by selecting specific cattle to sell either in the cash market or through alternative marketing agreements. High-quality cattle are more likely to be sold in the alternative market as prices are not fully calculated until after harvest, allowing carcass quality premiums to be added. Consequently, it is assumed that low-quality cattle are sold in the cash market to avoid potential carcass discounts. Depending on a feedlot's size, relationship with packers, and marketing costs, these selection assumptions may not be accurate and bias prices.

Key words: cattle quality, feedlot, generalized Roy model, Heckman model

Introduction

In the cattle procurement market, feeders can sell their cattle to packers through either alternative marketing agreements (AMA) or the cash market. Some researchers and producers suggest that packers negatively influence prices through either oligopsony power or an overuse of AMAs (Ward, 1999). With so much emphasis on packer power and quantity of transactions, little has been done to study the impact that feedlots and their marketing choices may have on fed cattle prices. Specifically, feeders can select which market to send a specific lot of cattle to, dictating the distribution of cattle quality between cash and AMA markets. The difference in cattle quality between these two markets has not been previously studied despite being one of the key components of fed cattle price formation (Koontz, 2010).

The AMA market, sometimes known as the captive supply market, refers to cattle that are committed to a certain buyer at least 2 weeks in advance of slaughter. AMA sales comprise three categories: packer owned and fed, forward contracts, and formula-based agreements such as grid-based pricing or futures-based pricing (Ward, Koontz, and Schroeder, 1998; Ward, 1999; Xia and Sexton, 2004; Ward, Schroeder, and Feuz, 2017; Xia, Crespi, and Dhuyvetter, 2019; Dennis). The cash market comprises all other live market sales and includes auction barns, dealers and brokers, and direct trade between packers and feeders. Although the two markets are separate, AMAs and cash sales are connected through price formation. Cattle that are procured through AMAs are ultimately priced using either the weekly regional cash price or the plant's average cash price as the base to which premiums and discounts are applied (Schroeter and Azzam, 2003; Zhang and Brorsen, 2010; Adjemian et al., 2016; Peel et al., 2020). If the cash sales were to be inaccurately priced at a lower value than what should be observed based on the quality of cattle found in this market, then past studies may be biased in their findings about market power as they assumed quality to

Kayla Hildebrand (corresponding author, kayla.hildebrand@oksate.edu) is a former graduate student and Chanjin Chung is a professor and the Breedlove Endowed Professor in the Department of Agricultural Economics at Oklahoma State University.
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.
Review coordinated by Jeffrey J. Reimer.

be consistent across all sales and markets during estimation. As a result of ignoring self-selection, market power parameters may be overestimated.[1]

Contrary to popular opinion, feeders hold the power over packers when it comes to dictating to which market a pen of cattle will be sent (Koontz, 2015). Packers may choose feedlots to conduct business with, but feedlots optimally decide which cattle to market to the packers by strategically grouping high-quality and more uniform pens together. Feeders also hold an information advantage over packers in the form of knowing the origin of the cattle, genetics, medical records, and other characteristics of their animals. Feeders can therefore better predict carcass quality and performance compared to packers and self-select which market to send cattle.[2] Given this information advantage, we hypothesize that feeders attempt to maximize profits by making strategic decisions about how to group cattle lots as well as how to market these cattle lots. This strategy may lead to selectivity bias, which may ultimately cause lower prices to be observed in the cash market. Specifically, packers assume feeders will send their high-quality cattle to the AMA market to capture price premiums. This assumption consequently suggests that low-quality cattle will be sent to the cash market, prompting packers to establish a cash base price based on low-quality stock. However, this may or may not be true for smaller feeders as they are more likely to solely utilize the cash market because of added independence and flexibility (Koontz, 2010); the quality distribution in the cash market may be unknown.[3] As a result, the cash market price may not reflect an accurate assessment of cattle quality and consequently may establish a lower base price than should be observed.

The objectives of this study are to test the presence of selectivity bias in the cattle procurement market and determine the impact that self-selection has on cattle prices. Unlike previous studies, our paper uses a unique dataset from a single feedlot that details transaction prices and quantities along with cattle quality information. The feedlot's AMA transactions provide post-harvest prices that are directly based on known carcass attributes, which is typically only known between the feedlot and the packer and is not mandatorily reported to any third-party data collection agency. To our knowledge, such detailed information has not been used in a fed cattle study. With these rare transaction data, we investigate the potential impact that selectivity bias may have on cash and AMA prices using the Heckman two-step model and the generalized Roy model. We extend our study from a single feedlot to the regional level by using a regional dataset collected by a private industry organization.

Literature Review

In the past 2 decades, researchers have used various theoretical and empirical frameworks to evaluate price formation and discovery in the fed cattle industry. Most of the existing literature has emphasized either packers' oligopsony power (e.g., Xia and Sexton, 2004; Chung and Tostão, 2009; Zhang and Brorsen, 2010; Ji, Chung, and Lee, 2017; Xia, Crespi, and Dhuyvetter, 2019) or the thinning cash market (e.g., Schroeter and Azzam, 2003; Koontz, 2015; Peel et al., 2020)

[1] Specifically, these previous studies have assessed the extent of oligopsony power through either quantity or price based conjectural models (e.g., Cournot, Stackelberg, and Bernard models) without accounting for the selectivity problem caused by asymmetric information about cattle quality. Therefore, the previous papers' conclusions regarding packers influencing prices may not be attributed solely to market power but also to selectivity bias. Selectivity bias may arise when feedlots self-select which cattle lots to sell in each market, creating a sample selection problem. The sample selection problem arises due to feedlots' strategic behavior, mainly driven by their information advantage regarding cattle quality.

[2] While completing this research, we spoke with feedlot managers who confirmed the strong record keeping and knowledge regarding the cattle they fed out.

[3] Small feedlots (operations with a capacity of 1,000 head or less) account for 95% of operations and 15% of the fed cattle market each year (US Department of Agriculture, 2022). As pointed out by a reviewer, smaller feedlots may not be able to compete and bid on high-quality feeder cattle to grow and finish out. As a result, they would theoretically be priced out of the market or unmotivated by the expected low prices in the cash market. However, if small feedlots can obtain lower priced feeder cattle (due to either purchasing riskier, lower quality calves or by feeding out their own stock from their cow-calf operation), then a small profit may still be realized. However, this is an important concern as the market shifts toward a smaller number of larger, more specialized feedlots.

without controlling for quality attributes. Arguments have ranged from claiming that the elimination of AMAs would decrease the effect of market power on prices to suggesting that a specified volume of cash sales is needed to ensure accurate price discovery (Koontz, 2010, 2015). Sabasi et al. (2013) suggest that if AMAs were widely used, then competition between the two markets would be reduced and would depress not only cash market prices but also prices from all procurement methods.

However, quality—although often left out of such analyses—plays an equally important role alongside supply and demand characteristics in price formation (Koontz, 2010). An early study on grid pricing by McDonald and Schroeder (2003) examines how cattle quality and feeding performance influence profit per head; premiums and discounts correlated to these factors have the potential to increase variability in selling price, where the study ultimately found that feeder cattle price and grid price (with consideration of cattle quality) have the greatest impact on profit per head and offer the largest opportunity to manage profit risk. Similarly, Anderson and Zeuli (2001) argue grid variability combined with pen quality differentials and carcass quality discounts leads to a marketing risk. This variability in price may be traced back to the overall cattle quality that is perceived in the cash market as well as to a specific AMA lot's quality. Establishing a base price, typically from either the average cash price or the plant's average price, may be the greatest concern regarding AMA pricing due to price discovery and pricing accuracy. These base prices may not represent the same quality of cattle being marketed on the grid and may consequently cause formula- and grid-based prices to decline (Ward, Schroeder, and Feuz, 2017). Peel et al. (2020) argues that if market information were symmetric and cattle quality were fully known between cash and AMA markets, then packers would be able to randomly select cattle in the cash market to establish a base price to produce a more accurate average price.

We argue that the nonrandom distribution of cattle quality between AMA and cash markets could potentially lead to the sample selection (or self-selection) problem, particularly in the cash market. Consequently, this could play a significant role in forming biased prices in the cattle procurement market.[4] To our knowledge, fed cattle prices have not been analyzed in the literature from the sample selection perspective. Although previous studies acknowledge that cattle quality in the cash market may negatively impact AMA cattle prices by establishing a lower base price, the actual impact that feeders' market selections have on prices has not been evaluated. A feedlot's strategic decision to market low-quality cattle in the cash market (to avoid price discounts in the AMA market) may still hinder the prices obtained from formula or grid-pricing. This is because the cattle population in the cash market is unrepresentative of the general market's overall quality. Self-selection models can allow one to determine whether the distribution of cattle quality between the cash and AMA market leads to a sample selection issue and how significant that issue may be regarding its impact on observed fed cattle prices.

A similar topic of adverse selection has been explored in the thoroughbred industry (Chezum and Wimmer, 1997; Wimmer and Chezum, 2003, 2006). Akerlof (1970) was the first to introduce the theory of the adverse selection problem regarding information asymmetry, where a seller knows the true quality of goods but the buyer only knows the distribution of the quality of goods. Wimmer and Chezum (2003) implement the Heckman sample selection model to analyze how adverse selection impacts the market in which racehorses are sold depending on horse quality, information asymmetry, and seller characteristics. Like the two correlated markets for fed cattle, the thoroughbred market has a similar setup: The base price in the certified market is established from the expected quality of horses sold in the auction market. As a result, breeders are expected to only sell their low-quality horses in the auction market and keep their higher quality horses for breeding purposes to avoid receiving low prices. The study found negative selection within the noncertified market, where the overall market was characterized by lower quality horses; the certified market did not present selectivity bias in prices as third-party certification alleviated most information asymmetries and allowed for more accurate pricing.

[4] Peel et al. (2020, p. 6) support this argument, claiming "differences in the type or quality of cattle traded by negotiated cash arrangement compared to formula cattle could be evidence of adverse effects of thin markets."

The main contributions of our study include the following. First, two proprietary datasets collected from both a feedlot and a regional industry organization are used for our study. As stated earlier, these unique datasets with detailed transaction-level information (particularly from the feedlot data) have not been used in fed cattle studies. Second, our paper studies the price-suppression problem in the cattle procurement market using the sample selection framework, while previous studies mostly attribute the same issue to the market power imbalance between feeders and packers caused by the packer concentration. Finally, we implement a well-known econometric procedure, the sample selection model, for the cattle procurement market. Although past studies have applied similar models to the thoroughbred industry, no such models have been applied to study the issue of potentially biased prices in the cattle procurement market.

Data

This study uses two unique fed cattle datasets. The first dataset covers transactions of a single feedlot in Oklahoma from November 2018 to July 2019. The feedlot has a 32,000-head capacity with facilities that can hold 60–300 head per pen.[5] The data include 398 lots of cattle with head counts of 1 to 212.[6] Carcass and price data cover a total of 18,097 head of cattle. The second dataset is proprietary information from an industry organization and includes weekly aggregate sales for an unnamed region between 2013 and 2019. These transactions occur over two different subregions for a total of 3,870 transactions and 1,073,078 head of cattle.

Dataset from a Feedlot

Table 1 reports summary statistics of key variables from the single feedlot dataset. Approximately 6.92% head of cattle were sold through the cash market, consistent with the 4%–10% cash trade typically found in the Oklahoma–Texas–New Mexico region (Koontz, 2015). The average cash market price was \$66.55/cwt at a minimum price of \$10.41/cwt and a maximum price of \$136.82/cwt. We acknowledge that our cash prices are lower than the average prices reported by the USDA for the same period (US Department of Agriculture, 2021).[7] We attribute this to the low head count of some lots (the average price for the 128 single-headed lots in our dataset is \$58.66/cwt), whereas the cash lots that contain 15 head or more have much higher average prices that are more in line with the typical USDA average (the average price for the 20 lots that had 15 head or more is \$114.80/cwt). If we were to delete lots with less than 15 head for the analysis, then we would be ignoring feedlot's strategically self-selecting not only how to market lots of cattle but also how to market single head of cattle to maximize profits. The remaining lots were AMA transactions: 74 lots were sold through the negotiated grid, 161 lots were sold through the US Premium Beef (USPB) grid, and 1 lot was sold via formula pricing. Cattle sold through AMAs received \$122.45/cwt on average, with a minimum price of \$99.32/cwt and a maximum price of \$134.70/cwt. The average AMA price is notably higher than the average cash price. This clear price difference produces a large impact on calculated profits, further providing evidence of a quality difference, or at least a quality perception difference, between the two markets that may be explained by this specific feedlot strategically selecting how to market their cattle to optimize profits.

---

[5] Due to the confidentiality agreement with the feedlot, the name of data sources cannot be disclosed. We obtained the data by traveling to the feedlot on two separate occasions and transferring printed records into our own data sheet.

[6] Certain lots were tactically split up and sold over different periods. Therefore, we report a minimum head per lot of one, where the single head of cattle was part of a larger lot of cattle but was sold either at a different time or through a different marketing method to maximize profits. Single-headed lots may be considered “cull” cattle and will only be marketed through the cash market, such as a livestock market, due to convenience, risk aversion tactics, and being unable to meet lot-size quotas demanded by packers.

[7] Between November 2018 and July 2019, USDA monthly feeder cattle prices ranged from \$113/cwt to \$126/cwt (US Department of Agriculture, 2021).

Table 1. Summary Statistics of Variables from a Single Feedlot Dataset

	Cash Sales (N = 163 lots)				AMA Sales (N = 236 lots)			
Variables	**Mean**	**S.D.**	**Min.**	**Max.**	**Mean**	**S.D.**	**Min.**	**Max.**
Lot price ($/cwt)	66.55	37.13	10.41	136.82	122.45	7.16	99.32	134.70
Head count	8	19.97	1	12	7	38.13	6	212
Pay weight (lb)	1,052	293.62	510	1,576	1,321	122.74	1,061	1,625
Avg. daily gain (lb)	2.03	1.88	−1.64	3.93	3.08	0.54	0.08	4.56
Gender								
Heifer ($)	63.86	35.82	10.41	126.00	123.22	7.43	99.32	133.73
Steer ($)	68.93	36.86	15.40	136.82	121.44	6.70	107.04	134.70
Mixed ($)	62.29	44.71	11.96	121.17	124.67	7.97	131.04	108.82
Buyer								
National Beef ($)	119.36	5.44	110.75	124.43	122.50	7.14	99.32	134.70
Tyson ($)	129.08	5.21	125.50	136.82	–	–	–	–
Cargill ($)	111.33	0.84	110.73	111.92	111.78	0.00	111.78	111.78
Other ($)	58.58	32.87	10.41	132.41	–	–	–	–
Location								
Liberal, KS ($)	103.10	30.48	30.90	124.36	121.95	7.34	107.04	133.73
Dodge City, KS ($)	118.73	6.48	110.73	124.43	124.00	6.38	99.32	134.70
Holcomb, KS ($)	129.08	5.21	125.5	136.82	–	–	–	–
Unknown ($)	59.10	33.30	10.41	132.41	–	–	–	–
Other								
Base live price ($/cwt)	–	–	–	–	120.31	6.46	109.13	129.90
Weekly avg. price ($/cwt)	122.2	5.23	108.91	129.77	119.82	6.43	108.88	129.90
Negotiated grid ($)	–	–	–	–	120.95	7.70	107.04	133.65
USPB base grid ($)	–	–	–	–	123.23	6.82	99.32	134.70
Formula ($)	–	–	–	–	111.78	0.00	111.78	111.78
Premiums and discounts								
Quality adjustment ($)	–	–	–	–	1.35	1.59	−4.96	8.62
Yield adjustment ($)	–	–	–	–	0.39	4.39	−5.43	3.18

Other variables include *head count*, *pay weight*, *avg. daily gain*, *gender*, *packer*, *location of plant*, *base live price* (i.e., the starting price to which premiums and discounts are applied), *weekly avg. price* of region, type of AMA, and carcass-specific variables reported after AMA cattle were slaughtered. Arguably the most unique aspect of this dataset is the post-slaughter variables, *quality adjustment* and *yield adjustment*. Quality is broken down into six grades: prime, choice or higher, select, no roll, hard bone, and dark cutter. Yield grade adjustment is broken down by numerical yield grade 1–5, where 1 corresponds to a carcass with a thin layer of fat and 5 corresponds to a carcass covered in extensive fat.[8] These two carcass measures, which—to the best of our knowledge—have not previously been used in any cattle procurement price study, should allow us to effectively control for cattle quality.

[8] At the time of slaughter, each head of cattle receives its own individual quality and yield grade. The lot's total premiums and/or discounts are then based off the cumulative performance of all cattle; a single, poor-performing carcass can negatively influence the entire lot's price within the AMA market. Such carcass information is not known at the time of a cash sale as animals are harvested after a negotiated price has been established.

Table 2. Summary Statistics of Variables from Cash Market Sales of an Industry Organization (*N* = 3,870 lots)

Variables	Mean	S.D.	Min.	Max.
Lot price ($/cwt)	122.10	12.04	94.00	172.00
Head count	299.00	297.98	5.00	2,781
2013 avg. price ($)	125.87	3.68	119.00	134.00
2014 avg. price ($)	149.55	8.19	139.00	172.00
2015 avg. price ($)	143.15	13.18	135.00	160.00
2016 avg. price ($)	119.03	11.58	97.00	140.00
2017 avg. price ($)	122.05	8.85	105.00	145.00
2018 avg. price ($)	117.13	6.59	99.00	130.00
2019 avg. price ($)	119.97	6.80	97.00	128.00
Gender				
Heifer ($)	122.38	11.78	97.00	172.00
Steer ($)	123.72	11.85	98.00	172.00
Holstein ($)	106.33	11.21	94.00	132.00
Buyer				
JBS ($)	123.18	13.52	98.00	172.00
Tyson ($)	123.17	9.92	104.00	164.00
Cargill ($)	122.17	11.24	97.00	166.00
National Beef ($)	124.91	11.88	98.00	170.00
Other ($)	120.79	14.01	94.00	162.00
Region				
Region A ($)	123.07	11.90	94.00	172.00
Region B ($)	123.10	11.77	97.00	172.00

Dataset from an Industry Organization

Table 2 reports summary statistics from an industry organization dataset. Although this specific organization reports that AMA sales account for 79%–100% of all cattle trades on a weekly basis since January 2019, it does not include cattle sale prices for AMA sales. As a result, the regional feedlot dataset is used solely for cash market analysis. Observations are on a weekly basis over a 7-year period and include information regarding *head count*, *gender*, *packer*, and *region*; therefore, a feedlot may have multiple observations in any given week, depending on whether cattle differ in gender or are purchased through a different procurement method. The average cash price received is $122.10/cwt, with a minimum price of $94.00/cwt and a maximum price of $172.00/cwt.

Cash market sales during this period comprised 2,089 lots of steers and 1,703 lots of heifers. *Cargill*, *JBS*, *Tyson*, *National Beef*, and *Other* purchased 1,618, 846, 500, 827, and 79 lots, respectively, from the cash market. A total of 2,666 lots were from subregion “A” and the other 1,204 lots were from subregion “B”. Due to the confidentiality of our regional data source, we are unable to disclose the different subregions. It is important, however, to include each subregion in analysis as Mallory et al. (2016) find that area-specific characteristics affect cattle prices.

Model

To identify the presence of selectivity bias in the cattle procurement market and determine its impact on cattle prices, we estimate two sample selection models: the Heckman model and the generalized Roy model. We argue that when feedlots select cattle they sell, the sample of cattle sold is censored strategically, which results in sample selection bias. Selectivity bias is identified by testing the coefficient of the sample selection variable, the inverse Mills ratio, and its impact is determined by examining whether observed prices are lower or higher than the prices generated from a random selection process (see similar approaches in Chiappori and Salanie, 2000; Wimmer and Chezum, 2003).

The starting point of theoretical and conceptual arguments of sample (i.e., self-) selection models is that rational actors make rational decisions about which markets to participate in under assumptions that there are different distributions of product quality across markets (assumption 1) and that the product quality of each market is correlated (assumption 2).[9] We argue that different distributions of cattle quality between AMA and cash markets are a direct result of a feeder's attempt to maximize profits through self-selection and that cattle quality between the two markets are highly correlated.

In the cattle procurement market, while assumptions (1) and (2) are fully satisfied, it is likely that feeders hold an information advantage with respect to cattle quality based on their knowledge of cattle genetics, origin, medical records, and other quality-related information. With an information advantage, feedlots optimally select and send their perceived high-quality cattle to the captive supply market and their remaining cattle to the cash market to best optimize their profits and avoid the risk of receiving discounts in AMA markets. For example, larger feedlots may strategically split up cattle lots and sell the perceived low-quality cattle to the cash market to keep one or more heads of cattle from negatively influencing a pen's overall price. Smaller feedlots and businesses, however, often use the cash market exclusively because of added independence and flexibility with regards to timelines, planning, and meeting demand quotas (Koontz, 2010). In this case, it is likely that smaller businesses will send both high- and low-quality cattle to the cash market. Packers may overlook this quality discrepancy among cattle supplied to the cash market from small and large feedlots. This discrepancy may lead to an unknown quality distribution in the cash market and as such, a selectivity bias impact on prices.

Specifically, a seller will market their cattle based on perceived quality attributes and how those quality attributes align with the pricing structure of each market j, $j \in \{1,2\}$, where 1 and 2 denote AMA and cash markets, respectively. Therefore, an individual feedlot i chooses the market j that maximizes its return,[10] Y_i:

$$Y_i = \max\left[Y_1^i, Y_2^i\right]. \tag{1}$$

We assume that each feedlot has two types of nonnegative cattle quality, Q_j, with corresponding positive prices π_j. The uppercase letter Q represents a random variable while the lowercase letter q represents its realization. We also assume that the quality Q_j is market specific and choice decisions are made based on feeder's information about cattle quality and price. From equation (1), a feedlot chooses market 1 over market 2 if the expected return from market 1 is greater (i.e., $Y_1 = \pi_1 Q_1 > Y_2 = \pi_2 Q_2$); otherwise, it chooses market 2.

Next, it is assumed that there exists a population distribution of quality, (Q_1, Q_2), and its well-defined density function, $f(q_1, q_2)$, and that for each pen of cattle a feedlot has a market preference for market 1 or 2, with no indifference between the two options. Following Heckman and Honoré

[9] Earlier studies also looked at different distribution of skills in the labor market (Autor, 2009).

[10] The feedlot chooses a market for each pen if the feedlot has more than one pen.

(1990), the proportion of fed cattle sent to market 1, M_1, can be stated as[11]

$$M_1 = \int_0^{\infty} \int_0^{\pi_1 q_1/\pi_2} f\,(q_1,q_2)\,dq_2 dq_1. \tag{2}$$

Equation (2) indicates that feeders would want to increase their cattle sales to market 1 as long as the relative prices related to quality attributes that they receive from market 1 are higher than those they receive from market 2. However, if more cattle are sent to market 1, the density of quality within this market differs from the overall fed cattle density of quality. Specifically, while the population density of cattle quality is[12]

$$f_1\,(q_1) = \int_0^{\infty} f\,(q_1,q_2)\,dq_2, \tag{3}$$

the quality density for market 1 is

$$g_1\,(q_1 \mid \pi_1 q_1 > \pi_2 q_2) = \frac{1}{M_1} \int_0^{\pi_1 q_1/\pi_2} f\,(q_1,q_2)\,dq_2. \tag{4}$$

Equation (4) shows that as feedlots make their optimal decisions, the distribution of quality attributes in market 1 differs from the general fed cattle population unless (i) the relative price of cattle quality attributes, π_1/π_2, becomes large, and as a result, M_1 becomes 1, where all cattle are sold in market 1 or (ii) $pi_1 q_1 > \pi_2 q_2$ is independent of q_1. Comparing equations (3) and (4) explains the selectivity problem stated in Roy (1951). The selectivity problem represented by the difference between the population density function in equation (3) and the conditional density function in equation (4) is expected to affect cattle quality and price from AMA and cash markets (and, accordingly, feeders' return from each market) differently.

With each pen of cattle differing in quality, feedlots will strategically shift between markets based on the expected profits that may be obtained for each lot. Strategic shifts between markets should alter the distribution of cattle quality and therefore cattle prices between the two markets as well as the overall cattle market. Sample selection models used in this study test for the strategic censoring nature of sample selection between AMA and cash markets and measure how the distribution of cattle quality between the two markets impacts the distribution of prices (see Borjas, 1987; Heckman and Honoré, 1990; Autor, 2009, for further discussions on conceptual and theoretical issues of sample selection models).[13]

The Heckman two-step correction method (type 2 Tobit) is arguably the most utilized procedure to study sample selection (Winship and Mare, 1992; Bushway, Johnson, and Slocum, 2007). Systematic censoring occurs when producers strategically select the animals they sell in each market (Wimmer and Chezum, 2003). This systematic selection is modeled in our study as a case of selectivity bias using the Heckman two-step procedure while allowing correlation between errors of selection and price equations. However, one can argue that the Heckman two-step procedure is appropriate to characterize incidental selection that occurs when truncation is applied to a stochastic function of the dependent variable where the observed predictor variable is not the actual selection variable but is correlated with it (Berk, 1983; McGuire, 1986; Winship and Mare, 1992; Bushway, Johnson, and Slocum, 2007). Self-selection may more accurately describe the fed cattle

---

[11] Choosing any $S = \sim 2$, the probability of $(Q_1, Q_2) \in S$ must be 1. Therefore, $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(q_1, q_2)dq_2 d1_1 = 1$. However, the lower bounds of both integrals should be 0 due to the nonnegativity assumption on cattle quality, Q. We also impose the feeder's selection rule for market 1, $0 < q_2 < \pi q_1/\pi_2$, derived from $\pi_1 q_1 > \pi_2 q_2$.

[12] In this case, the quality density functions from markets 1 and 2 have the same form. That is, $f_1(q_1) = \int_0^{\infty} f(q_1, q_2)dq_2 = \int_0^{\infty} f(q_1, q_2 dq_1 = f_2(q_2)$.

[13] Some researchers may recommend conducting a propensity score matching study versus sample selection. However, propensity score matching's conditional independence assumption limits the procedure from allowing the estimated average treatment effect from being subjected to unobserved selection bias (Peel and Makepeace, 2012).

market, where producers and feeders make rational, optimizing decisions regarding which markets to participate (Autor, 2009). Therefore, as an alternative, we also estimate the generalized Roy model (type 5 Tobit) as it better captures self-selection by estimating the correlation between the selection equation and the two price equations simultaneously.

The theory behind the generalized Roy model originates from an example of optimizing wealth in the labor market, where Roy hypothesized workers given varying skillsets may self-select into a specific sector to participate depending on expected comparative advantages (Roy, 1951; Heckman and Honoré, 1990; Autor, 2009). The theory was then transformed by Borjas (1987) into a switching-regression model to analyze how immigrant workers may decide in which country to work based on expected wages. The generalized Roy model assumes the individual compares opportunity costs to make an optimal choice, which is more in line with the decision process feeders face when choosing how to market a lot of cattle to maximize their profit.

Heckman Two-Step Model

The Heckman two-step procedure estimates the effect that feedlots' market selection has on observed prices that is characterized by selectivity bias. The selectivity bias can cause observed prices to be lower than prices that may have been formed through a random selection process. Feeders can sell their cattle in either the cash market or the captive supply market. Given this dichotomous choice and the assumption that feeders select the market they believe will maximize profit for a specific lot of cattle, our analysis begins with the following probit model:

$$z_i^* = \boldsymbol{w}_i'\boldsymbol{\gamma} + \mu_i, z_i = \begin{cases} 1 & \text{if } z_i^* > 0 \\ 0 & \text{if } z_i^* \leq 0 \end{cases}, \tag{5}$$

where z_i^* is a latent variable associated with the choice of market; z_i denotes the feeder's binary choice of pricing method ($z_i = 1$ corresponds to cattle sold in the cash market and $z_i = 0$ corresponds to cattle sold in the AMA market); i corresponds to the lot of cattle that is sold; $\boldsymbol{w}_i'\boldsymbol{\gamma}$ is a vector of specific attributes related to factors that may influence a feeder's market selection and their corresponding estimated coefficients; and μ_i is the normally distributed error term with $\mu_i \sim N i.i.d.(0,\sigma_\mu^2)$.

The second step of the Heckman procedure involves the estimation of price equation, where the inverse Mills ratio (IMR) is included to account for incidental truncation caused from feeders' selection process. The IMR, also known as a hazard rate measure, is calculated from the variables predicted in equation (5):

$$\lambda\left(\boldsymbol{w}_i'\boldsymbol{\gamma}\right) = \frac{\phi\left(\boldsymbol{w}_i'\gamma\right)}{\Phi\left(\boldsymbol{w}_i'\gamma\right)}, \tag{6}$$

where $\phi(\boldsymbol{w}_i'\boldsymbol{\gamma})$ and $\Phi(\boldsymbol{w}_i'\boldsymbol{\gamma})$ are the probability density function and cumulative distribution function, respectively (Heckman, 1979). Transaction price, y_i, is specified as

$$y_i = \boldsymbol{\alpha}'\boldsymbol{X_i} + \rho\sigma\lambda\left(\boldsymbol{w}_i'\boldsymbol{\gamma}\right) + \varepsilon_i, \tag{7}$$

where $\boldsymbol{\alpha}'\boldsymbol{X}_i$ represents attributes of cattle lot i used for price formation and corresponding coefficients and ε_i is the normally distributed error term $\varepsilon_i \sim N i.i.d.(0,\sigma_{\varepsilon_i}^2)$. If cattle lots are self-selected into the cash market, then unobservable factors (e.g., post-harvest carcass quality characteristics) that decrease (increase) cattle prices increase (decrease) the probability that a lot is sold in this specific market. The nonrandomness in choosing a market could lead to selection bias without incorporating IMR, $\lambda(\boldsymbol{w}_i'\boldsymbol{\gamma})$ in equation (7). The selection bias is captured by the coefficient of IMR, where ρ represents the correlation between the error terms of selection equation (5) and price equation (7) and is calculated as

$$\rho = \frac{\sigma_{z,y_i}}{\sigma_{z}, \sigma_{i}}, \tag{8}$$

where σ_{z,y_i} is the covariance between the error terms of the selection and price equations; σ_z and σ_{y_i} are the square roots of error variances from selection equation (5) and price equation (7), respectively; and σ is the standard error from the price equation. Selection effects are captured in $\rho\sigma$, which represent the impact of IMR on the variation of cattle prices. A large and significant IMR value suggests there is a high probability that the market (with $z = 1$) is chosen. A negative and significant IMR coefficient indicates unobservable factors that would increase the probability of a cattle lot sold in the market with $z = 1$ (cash market in our case) is inversely correlated with unobservable factors that would increase prices. A positive and significant IMR coefficient corresponds to the same unobservable characteristics being positively correlated with unobservable factors that would increase price. An insignificant IMR coefficient indicates no selection bias.

The traditional Heckman model includes a single regression equation in the second stage, solely focusing on data points that are above a certain cut-off point. Our research problem consists of two price regressions: the cash market price and the AMA market price equations. The current IMR is based on the probability that the cash market is selected and needs to be re-solved with respect to the probability that a lot of cattle is selected to be sold in the AMA market. We hypothesize that premiums and discounts applied in the AMA market alleviate information asymmetries and consequently allow for more accurate pricing. Specifically, cattlemen are incentivized by premiums to commit their perceived high-quality lots to the captive supply and are punished by discounts for sending cattle that are below average quality. Therefore, we expect that the IMR coefficient becomes statistically either less significant or insignificant, which would indicate less or no selection bias in the captive supply market.

Generalized Roy Model

Alongside the Heckman model, we also estimate the generalized Roy model as an alternative method to test for selectivity bias. The generalized Roy model begins the same as the Heckman model: the IMR is calculated from equation (1). However, in the second step, price equations, y_1 (when $z = 1$) and y_2 (when $z = 0$), are estimated simultaneously for cash and captive supply markets. Although only one price may be observed for each sale, it is important to remember that the two price equations are connected to each other because the cash price serves as the base price for the AMA price. When interpreting the IMR coefficient, a positive and significant $\rho\sigma$ signifies positive selection, where cattle lots are positively selected into a market and are above the average of that market's quality distribution. Negative selection is characterized by cattle lots selected for a market having an average quality distribution below that of the overall market. Negative selection would ultimately suggest that low-quality cattle are sent to the cash market in an attempt to try to garner higher prices without providing full quality information.

Effect of Selectivity Bias on Cattle Price

We estimate the extent to which selectivity bias may impact cattle prices following Lee's (1995) opportunity cost approach. Unlike Lee's paper, we continue with the probit selection model instead of switching to a logit for two reasons: to stay consistent with our two previous self-selection models and to properly estimate a hazard rate variable (Bushway, Johnson, and Slocum, 2007). To estimate the magnitude that selectivity bias has on prices, we first calculate the expected price that a random lot of cattle would receive if sold to the cash market, $E[P_{cash}|\alpha,X]$. To calculate this measure, the cash price regression results are used to calculate predicted cash prices for all lots of cattle in our

Table 3. Estimates of Selection Equations with Single Feedlot Data

	Heckman Model		Roy Model	
Variable	**Estimate**	**Std. Err.**	**Estimate**	**Std. Err.**
Intercept	5.40	3.37	6.42*	3.33
Head count	−0.31***	0.04	−0.33***	0.04
Avg. daily gain	−0.67***	0.21	−0.67***	0.20
Liberal	−1.71***	0.27	−1.65***	0.25
Steer	0.77***	0.25	0.72***	0.25
Weekly avg. price	−0.02	0.03	−0.03	0.03
Quarter 1[a]	0.33	0.43	0.56	0.43
Quarter 2[a]	0.46	0.40	0.39	0.40
Quarter 4[a]	0.96**	0.48	1.00**	0.48
Log likelihood	−74.62	–	−1,327	–
Akaike information criterion	167.25	–	2,719	–

Notes: Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% levels, respectively.
[a]Quarter 3 was dropped as the reference period.

dataset (i.e., both AMA and cash lots); the average of these new predicted prices is then taken to establish the expected price that a randomly selected lot would receive in the cash market. We follow a similar procedure to estimate the expected price that a lot of cattle would receive if the lot were sold in the cash market, $E[P_{cash}|\alpha,X,Z=1]$. To do this, we once again use the cash price regression results to calculate a predicted cash price for any lot that was sold in the cash market (i.e., exclude the AMA lots from this estimation) before averaging the predicted prices to estimate the expected cash price of a lot that was selected into the cash market. Opportunity costs are then calculated by taking the difference between the expected price of a randomly selected lot and the expected price of a lot sold through the cash market:

$$E\left[P_{cash} \mid \alpha,X\right] - E\left[P_{cash} \mid \alpha,X,Z=1\right]. \tag{9}$$

Results

All models are estimated using the SAS QLIM procedure, and corrected standard errors are also automatically calculated under the procedure (SAS Institute, Inc., 2014). Before estimating each market's price equations to examine selectivity bias, we first estimate the selection equation for both the Heckman model and the generalized Roy model. Table 3 reports coefficients of the selection equation estimated using the single feedlot data. Variables considered in this equation include *head count* (number of cattle from each lot), *avg. daily gain*, *liberal* (a dummy variable equal to 1 when the cattle lot is sent to a packer plant located in Liberal, Kansas, and 0 otherwise), *steer* (a dummy variable equal to 1 when the cattle lot is steers, and 0 otherwise), *weekly avg. price* (weekly average cattle price in the region in which the feedlot is located), *Quarter 1*, *Quarter 2*, and *Quarter 4* (quarterly dummy variables to represent the specific calendar quarter a given lot was sold; Quarter 3 serves as the reference). Estimating these selection models helps determine whether there exists a systematic difference in the decision of marketing cattle between cash market versus noncash market options. It is important to consider systematic differences between the two market options as the models may detect potential selection bias (Cuddeback et al., 2004).

Both the Heckman model and the generalized Roy model show similar estimates and standard errors, presenting only slight differences. The coefficient for *head count* is negative and significant at the 1% level for both models, suggesting that lots with a greater number of head of cattle are less likely to be selected into the cash market (selection = 1) than to be sent to the AMA market (selection = 0). The result is consistent with feeders strategically culling a low number of poorer

Table 4. Estimates of Price Equations for Cash Market Sales with Single Feedlot Data

	Heckman Model		Roy Model	
Variable	**Estimate**	**Std. Err.**	**Estimate**	**Std. Err.**
IMR variable				
$\rho\sigma$	−26.36**	11.99	−21.40***	4.89
Other variables				
Intercept	8.36	8.72	7.38	8.58
Head count	7.90***	2.61	6.59***	1.50
Liberal	20.74*	12.24	16.86*	9.13
Pay weight	6.96***	1.04	7.07***	1.02
Steer	−9.04*	4.89	−9.32**	4.71
Quarter 1[a]	2.94	7.19	4.11	7.08
Quarter 2[a]	13.79*	7.47	13.24*	7.31
Quarter 4[a]	−4.09	7.68	−3.66	7.43

Notes: Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% levels, respectively.
[a]Quarter 3 was dropped as the reference period.

quality cattle from their initial lot and selling in the cash market to avoid incurring discounts in the AMA market. The coefficient for *avg. daily gain* is also negative and significant at the 1% level, suggesting a lot that gains more efficiently is likely to be sent to the AMA market. This indicates that fleshier and easier weight-gaining cattle are better suited to be priced in the premium/discount market, whereas poorer-gaining or lighter-weight cattle are more likely sold in the cash market. The coefficient for *liberal* is also negative and significant at the 1% level, suggesting that cattle lots sold to the AMA market are more likely to be sent to a packer in Liberal, Kansas, whereas a cash lot is more likely to be sent to a packer elsewhere (i.e., Dodge City, Kansas; Holcomb, Kansas; or an unknown location). The coefficient of *steer* is positive and significant at the 1% level, which indicates that steers are more likely to be sent to the cash market compared to heifers or mixed lots of cattle. The estimate of *weekly avg. price* is negative but insignificant. *Quarter 4* is positive and significant at the 5% level, compared to Quarter 3, which suggests that cattle are more likely to be sent to the cash market over the AMA market in Quarter 4. Log-likelihood and Akaike information criterion (AIC) statistics indicate that the single feedlot data fit better in the Heckman two-step model than in the generalized Roy model.

Selectivity Bias in the Cash Market

Table 4 includes estimates from the Heckman and generalized Roy price equations; these equations are estimated from the single feedlot data. The existence of the self-selection issue is examined by testing the coefficient of IMR, $\rho\sigma$. The IMR corrects for selectivity bias by controlling factors that discriminate between selection of cash market and noncash market sales (Cuddeback et al., 2004). In our case, we test to see whether carcass quality characteristics and the disparate knowledge of such cattle characteristics between buyers and sellers may contribute to the selectivity bias. Under the Heckman model, the IMR coefficient is −26.362 and is significant at the 5% level. The IMR coefficient from the Roy model is −21.403, also significant at the 1% level. Negative and significant IMR estimates are consistent with the presence of selection bias (Wimmer and Chezum, 2006). Specifically, the lack of information during a sale leads to omitted variable bias that may influence both the probability of a cattle lot entering the cash market and the resulting price (Certo et al., 2016). Carcass information is revealed at different points during the transaction process for AMA and cash sales. During a cash transaction, packers can only take in account observable physical characteristics of the live cattle and any other limited knowledge they may have about the seller. Conversely, prices in the AMA market are established after the cattle are harvested to account for

Table 5. Estimate of Selection and Cash Price Equations with an Industry Organization Data

	Heckman Model	
Variable	**Estimate**	**Standard Error**
Selection equation		
Intercept	3.11***	0.28
Steer	−0.10	0.13
Head count	0.01	0.23
2013	−0.10	0.31
2014	−0.60**	0.30
2015	−0.40	0.41
2016	−0.64***	0.23
2018	−0.48**	0.24
2019	−0.36	0.25
Cargill	−0.51***	0.19
Tyson	0.18	0.36
JBS	−0.12	0.24
Cash price equation		
Intercept	115.77***	0.59
$\rho\sigma$	−40.62**	20.48
Steer	0.08*	0.36
2013	4.02***	0.63
2014	27.09***	0.99
2015	24.99***	1.12
2016	−0.41	0.89
2018	−3.54***	0.69
2019	−4.89***	0.61
Quarter 1	11.30***	0.32
Quarter 2	10.59***	0.31
Quarter 4	2.43***	0.29
Cargill	0.16	0.74
Tyson	1.07*	0.63
JBS	−1.06**	0.54

Notes: Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% levels, respectively. Dropped reference groups are 2017, National Beef and other packers, and Quarter 3.

premiums and discounts for traits like yield and quality performance. Therefore, packers can more accurately price the cattle in AMA sales because carcass performance is known. Also, packers are likely to have repeat contracts with feedlots and develop a relationship and ultimately a more in-depth understanding of the quality of cattle associated to that feedlot. However, feeders should have more information about cattle quality, both desirable and undesirable traits, particularly by cattle lot. Omitting carcass quality characteristics in the cash market's transactions leads packers to establish prices with limited information, which in turn negatively affects cash prices. Alongside limited information about cattle quality, cash prices may also be negatively impacted by other factors such as packers' risk aversion behavior.

Table 4 also reports estimates of *head count*, *pay weight*, *steer*, *liberal*, *Quarter 1*, *Quarter 2*, and *Quarter 4*. The coefficient on *head count* is positive and significant at the 5% level, indicating that lots with a greater number of cattle receive higher prices. The result makes intuitive sense, as single-headed lots have typically lower cattle quality. The *liberal* coefficient is positive and significant at the 10% level, suggesting that cattle sent to packers in this location receive higher prices compared to other packer locations (i.e., Dodge City, Kansas, and Holcomb, Kansas). The coefficient for *pay weight* is positive and significant at the 1% level, where heavier weight lots receive greater prices than lighter weight lots (which may be a result of light muscled livestock). The coefficient of *Quarter 2* is positive and significant, while estimates of *Quarter 1* and *Quarter 4* are insignificant at the 10% level.

To assess the validity of our findings from an individual feedlot, we also estimate the Heckman model using a different proprietary dataset collected from a regional industry organization. We are unable to estimate the generalized Roy model with this dataset as all AMA transactions have missing price information. Table 5 reports both selection and cash price equations' estimates. Results from the selection equation show that only a few estimates are statistically significant. Among dummy variables representing sales years, 2014, 2016, and 2018, are negative and significant when compared to the reference year 2017. Packers included in the regional analysis include *Cargill*, *Tyson*, and *JBS*, where *National Beef* (and *Other*) is dropped to serve as the reference class. Only *Cargill* is negative and significant at the 1% level, indicating that *Cargill* tends to rely less on the cash market (and more on the AMA market) to procure their cattle compared to *Cargill*.

The IMR coefficient from the regional data's cash price equation is −40.624 and is statistically significant at the 5% level. The result is consistent with our earlier findings from the single feedlot; a feedlot's ability to self-select the market to which they send a pen of cattle creates selectivity bias within the cash market due to an unknown quality distribution. Other variables considered in the cash price equation are *steer*, dummy variables representing sale years and quarters, and cattle buyers. *steer* is positive and significant at the 10% level, suggesting that steers receive higher cash prices compared to sale observations that are heifers or Holsteins (the dairy breed was included as a gender category in the regional dataset). The coefficients for dummy variables—*2013*, *2014*, and *2015*—are all positive and significant (compared to the reference year, 2017), while the coefficients for 2018 and 2019 are negative and significant. Quarterly dummy variables are all positive and significant at the 1% level. Dummy variables representing *Tyson* and *JBS* are also statistically significant at the 10% level.

Selection Bias in the AMA Market

We also examine whether selection bias is a concern in the AMA market. Only the single feedlot dataset is used for this analysis because, as indicated earlier, all AMA transactions have missing price information in the regional dataset. Table 6 reports regression results from the Heckman model and the generalized Roy model. Variables used for the AMA market's estimations are the same as those used for the cash market, with the addition of quality-related, post-harvest variables. We hypothesize that the post-harvest quality attributes used in premiums and discount pricing minimize any potential information asymmetries or selection bias regarding cattle quality. From the Heckman model, the coefficient of IMR, $\rho\sigma$, is −0.383 but insignificant. The generalized Roy model produced a corresponding coefficient value of −0.250. The results confirm our earlier hypothesis. The AMA market's premium/discount pricing system allows cattle buyers to generate more informed sale prices. As a result, our analysis finds that selectivity bias does not have a significant impact on the AMA market.

Estimates of *head count*, *liberal*, and quarterly dummy variables are all positive but insignificant. The coefficient of *pay weight* is negative and significant at the 1% level, which suggests that heavier weight lots do not necessarily receive higher prices in the AMA market after accounting for quality and yield adjustments. The *steer* coefficient is positive and significant at the 5% level. The variable *base live price* is the average price used to apply premiums and discounts, where the estimate is positive and significant at the 1% level, suggesting that a higher base price helps generate a higher AMA price. The estimate for *negotiated grid*, a dummy variable that corresponds to lots marketed through the negotiated grid (versus the USPB grid or formula marketing), is negative and significant at the 1% level, suggesting that these specific feedlots tend to receive lower prices than lots sold

Table 6. Estimate of Price Equations for AMA Market Sales with Single Feedlot Data

	Heckman Model		Roy Model	
Variable	**Estimate**	**Std. Err.**	**Estimate**	**Std. Err.**
IMR variable				
$\rho\sigma$	−0.38	0.52	−0.25	0.37
Other variables				
Intercept	2.58	4.2	3.35	3.85
Head count	0.13	0.14	0.04	0.04
Liberal, KS	0.11	0.82	−0.36	0.35
Pay weight	−1.13***	0.2	−1.20***	0.17
Steer	1.44**	0.57	1.70***	0.41
Base live price	1.10***	0.03	1.10***	0.03
Negotiated grid	−0.82***	0.29	−0.83***	0.29
Quality adjustment	0.88***	0.09	0.88***	0.09
Yield adjustment	−0.01	0.03	−0.01	0.30
Quarter 1[a]	0.66	0.65	0.76	0.62
Quarter 2[a]	0.42	0.54	0.59	0.49
Quarter 4[a]	0.76	0.85	1.12	0.68

Notes: Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% levels, respectively.

through the *USPB grid* and *formula-based pricing*. Quality grades are positively correlated to prices, where a lot that averaged a higher grade value will receive a greater monetary premium versus a pen of cattle that may have had a greater number of select (or lower grading) cattle. The positive and statistically significant coefficient of *quality adjustment* indicates that cattle sold through AMAs are high-quality and received added premiums. The variable *yield adjustment* corresponds to a lot's average yield grade based on fat depth. Yield grade is negatively correlated with price, where the negative and significant coefficient indicates that higher numerical yield grade cattle (i.e., cattle with excessive finish compared to frame size and muscle mass) receive lower AMA prices.

Effect of Self-Selection in the Cash Market

Peel et al. (2020) argue that if market information were symmetric and cattle quality were fully known between cash and AMA markets, then packers would be able to randomly select cattle in the cash market to establish a base price to produce a more accurate average price. However, if selectivity bias exists in the cash market due to the information asymmetry, then the base price may be distorted. To determine the impact of self-selection on the cash market, we calculate differences in expected prices of randomly selected lots and lots that were sold specifically in the cash market (Lee, 1995).

First, we calculate the predicted prices for each observation in our dataset using the results generated by the Heckman and Roy models' estimations. Then, we create four categories based on lot size (head count), where each lot size category consists of four subcategories to allow the comparison of randomly selected lots from cash lots for both the Heckman and Roy models. Randomly selected lots comprise all lot observations (i.e., both cash and AMA sales) and the cash lot subcategory comprises only lot observations that were marketed through the cash market. Expected prices are then calculated by averaging all the predicted prices that fit within each category. Next, expected total revenue is calculated by multiplying the expected price by average head count and average pay weight. To ensure a fair comparison between randomly selected lots and cash lots, we use the average head count and pay weight from the cash market for all revenue calculations in each category, respectively. Then, the degree of selectivity bias is calculated by taking the difference

Table 7. Impact of Adverse Selection on Cattle Procurement Market Prices

	Heckman Model			Roy Model		
	Randomly Selected Lots	Lots Sold in Cash Market	Difference from Adverse Selection Effect	Randomly Selected Lots	Lots Sold in Cash Market	Difference from Adverse Selection Effect
Lots with 1–5 head						
Expected cash price ($/cwt)	35.87	35.87	0.00	40.38	40.38	0.00
Expected total revenue ($/lot)	274	274	0.00	274	274	0.00
Average head count	1	1	–	1	1	–
Average pay weight	765	765	–	765	765	–
Lots with 6–30 head						
Expected cash price ($/cwt)	103.77	84.79	18.98	103.83	85.08	18.75
Expected total revenue ($/lot)	21,065	17,212	3,853	21,077	17,271	3,806
Average head count	20	20	–	20	20	–
Average pay weight	1,015	1,015	–	1,015	1,015	–
Lots with 31–60 head						
Expected cash price ($/cwt)	129.50	121.41	8.09	126.15	119.40	6.75
Expected total revenue ($/lot)	89,096	83,530	5,566	86,791	82,147	4,644
Average head count	50	50	–	50	50	–
Average pay weight	1,376	1,376	–	1,376	1,376	–
Lots with > 61 head						
Expected cash price ($/cwt)	156.60	146.25	10.35	166.09	152.76	13.33
Expected total revenue ($/lot)	192,057	179,364	12,693	203,696	187,348	16,348
Average head count	89	89	–	89	89	–
Average pay weight	1,378	1,378	–	1,378	1,378	–

Notes: Randomly selected lot calculation includes all lots (i.e., all AMA and cash transaction observations); values are based on the Heckman and Roy model estimations. Difference is calculated by subtracting the expected cash price from the expected random price. Expected total revenue (lot) is calculated by multiplying the expected price by average head count and pay weight. Average head count and pay weight are based off the cash observations within each category and used for both random and cash revenues to maintain consistency.

between random lots and cash lots. This calculation provides the potential impact that self-selection may have on fed cattle prices in the cash market.

Table 7 reports self-selection effects on cash prices. The first category does not include any AMA sales within the randomly selected lot category, as AMA sales traditionally commit and sell larger groups of cattle to packers. As a result, no difference in prices is observed. The remaining three categories, therefore, provide more insight on the potential monetary magnitude of selectivity bias. Expected price differences per hundredweight range from roughly $8–$19 (Heckman) and $7–$13 (Roy), translating into roughly a $3,850–$12,700 (Heckman) and $3,800–$16,350 (Roy) difference in expected revenue per lot. This substantially large price discrepancy suggests that cattle sold through the cash market are assumed to be of notably lower quality than cattle sold through the AMA market. As discussed previously, this quality perception may or may not be true, as quality may be skewed within the cash market as smaller businesses exclusively use this more convenient sale option versus creating deals with packers.[14] As a result, the randomly selected prices could more accurately portray the prices that should be observed based on an unknown quality distribution. Regardless, the comparison of expected prices and revenues provides a monetary measure to describe the selection bias impact in the cash market.[15]

Conclusion

Our study examines the potential self-selection problem in the cattle procurement market using two different sample selection models: the Heckman model and the generalized Roy model. The two models are applied to two unique proprietary transaction datasets collected from a feedlot and a regional industry organization. Then, using estimates of both models, we expand our analysis to simulate expected prices and revenues to demonstrate the potential monetary effect selectivity bias may have on cattle prices in the cash market.

Our results suggest that there exists a self-selection issue in the cash cattle procurement market. The self-selection is characterized by the negative selection caused by unobservable attributes (e.g., carcass quality), which are unknown until after the completion of the sale. Comparing the expected cash price from randomly selected lots of cattle and the expected cash price from cattle lots that were sold in the cash market show significant revenue loss to feeders.

Our results provide a few important implications that can help better understand current price discovery problems in the cattle procurement market. First, cash prices may have been determined at a lower level than they should be due to the selectivity bias caused by information asymmetry. The low cash prices lead to revenue loss to feeders, particularly small feeders, who do not have the volume to meet demand or establish relationships with large packers and who typically rely on the spot market for its convenience. Most feedlots in the United States are small operations, but these feedlots only generate about 15% of the fed cattle market each year (US Department of Agriculture, 2022). Second, the lower cash price also leads to lower overall cattle price, including those from the AMA market because the cash price is the basis for the AMA market price. Third, under the current price discovery system with its low cash prices, the cash market will be continuously depleted and thinned. The cash market is already thin; in some regions, the cash trade is less than 10%, which is not significant enough to provide an accurate base price for fed cattle. Last, our findings suggest that packers' market power exertion to lower cattle procurement price may have been over-estimated in the literature. Some of market power effect found in the literature could be attributed to the lower cash price caused by selectivity bias. Many earlier studies implemented quantity-based models that ignored quality differences across cattle lots (maybe due to restricted access to carcass-quality data) and feedlots' strategic behavior in marketing individual cattle lots.

The recently proposed Cattle Price Discovery and Transparency Act recognizes some of our findings stated above and requires minimums for negotiated (cash) sales and clear reporting of marketing contracts.[16] Some farmers might oppose the proposed act because they could be concerned

[14] One of our reviewers pointed out that small feeders may be unable to obtain high-quality stocker cattle because (i) there is little incentive to feed out if selling straight to the cash market, where lower prices are expected, and (ii) they may be outbid by larger operations in securing such cattle. Our paper does not specifically investigate the feeder cattle market, the distribution of cattle within the feeder cattle market, and how and from where smaller feeders procure cattle. Smaller businesses and operations likely obtain cattle from local livestock auctions and/or breed their own feeder cattle that they feed out as fed cattle. This specific topic is worth exploring further but is beyond the scope of our paper.

[15] Average cattle prices in the larger lot size categories surpass the average sale prices reported in Table 1. We attribute the difference to the numerous single head lots with significantly lower prices than cattle prices from larger lots.

[16] Specifically, if passed, the act would establish 5–7 regions of cattle markets in the United States and establish minimum levels of fed cattle purchases through approved pricing mechanisms and would penalize any large packer (i.e., any packer that has slaughtered 5% or more of cattle nationally in the past 5 years) that did not abide; it would also establish a publicly available library of mandated pricing reports and contract marketing in an effort to increase price transparency and competition (Henderson). The approved pricing mechanisms include fed cattle purchases made through negotiated cash, negotiated grid, at a stockyard, and through trading systems that multiple buyers and sellers regularly can make and accept bids.

about the possibility that it would limit cattle feeders' ability to use the AMA market. However, it should certainly increase the trade volume in the cash market, make the cattle procurement market more competitive, and—most importantly—help make the cash market work and put it back as the bedrock of the cattle pricing system. By strengthening the cash market, the proposed act could also help small operations relying mostly on the cash market.

Another way to improve price discovery in the cattle procurement market might be to improve the flow of market information on cattle quality using well-developed quality certification programs. The Oklahoma Quality Beef Assurance (OQBA) is one example of a certification program that leads feeder cattle producers to practice productive and ethical management techniques to enhance their herd's performance and profitability. Expanding a similar program to the feeder-packer level might help alleviate information asymmetries and consequently help establish more accurate prices based on cattle quality regardless of how the cattle are marketed.

Without knowing cattle prices associated with AMA sales in our industry organization data, our AMA analysis is restricted to prices reported by a single feedlot. This consequently limits our findings for AMA sales to a single feedlot versus a larger population of feeders. Therefore, it is recommended that future studies include multiple feedlots and regions as different regions offer different AMA methods as well as have drastically different trade ratios between cash and AMA sales. Given the significance of our findings, future studies are also recommended to consider effects of cattle quality on cattle price discovery, packers' market power, and other government policy issues.

[First submitted October 2021; accepted for publication July 2022.]

References

Adjemian, M., B. W. Brorsen, W. Hahn, T. L. Saitone, and R. J. Sexton. 2016. *Thinning Markets in U.S. Agriculture*. Economic Information Bulletin 148. Washington, DC: USDA Economic Research Service.

Akerlof, G. A. 1970. "The Market for "Lemons": Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84(3):488–500. doi: 10.2307/1879431.

Anderson, J. D., and K. A. Zeuli. 2001. "The Revenue Risk of Value-Based Pricing for Fed Cattle: A Simulation of Grid vs. Average Pricing." *International Food and Agribusiness Management Review* 4(3):275–286. doi: 10.1016/S1096-7508(01)00065-9.

Autor, D. H. 2009. *Self-Selection—The Roy Model*. Lecture Note 5. Massachusetts Institute of Technology: Cambridge, MA. Available online at https://web.archive.org/web/20200125142731/https://economics.mit.edu/files/15380.

Berk, R. A. 1983. "An Introduction to Sample Selection Bias in Sociological Data." *American Sociological Review* 48(3):386–398. doi: 10.2307/2095230.

Borjas, G. J. 1987. "Self-Selection and the Earnings of Immigrants." *American Economic Review* 77(4):531–553.

Bushway, S., B. D. Johnson, and L. A. Slocum. 2007. "Is the Magic Still There? The Use of the Heckman Two-Step Correction for Selection Bias in Criminology." *Journal of Quantitative Criminology* 23(2):151–178. doi: 10.1007/s10940-007-9024-4.

Certo, S. T., J. R. Busenbark, H.-s. Woo, and M. Semadeni. 2016. "Sample Selection Bias and Heckman Models in Strategic Management Research." *Strategic Management Journal* 37(13): 2639–2657. doi: 10.1002/smj.2475.

Chezum, B., and B. Wimmer. 1997. "Roses or Lemons: Adverse Selection in the Market for Thoroughbred Yearlings." *Review of Economics and Statistics* 79(3):521–526. doi: 10.1162/rest.1997.79.3.521.

Chiappori, P., and B. Salanie. 2000. "Testing for Asymmetric Information in Insurance Markets." *Journal of Political Economy* 108(1):56–78. doi: 10.1086/262111.

Chung, C., and E. Tostão. 2009. "Nonparametric Estimation of Oligopsony Power in First-Price Auction." *Journal of Agricultural Economics* 60(2):318–333. doi: 10.1111/j.1477-9552.2008.00188.x.

Cuddeback, G., E. Wilson, J. G. Orme, and T. Combs-Orme. 2004. "Detecting and Statistically Correcting Sample Selection Bias." *Journal of Social Service Research* 30(3):19–33. doi: 10.1300/J079v30n03_02.

Dennis, E. 2020, June 10. "Changing Grid Premiums and Discounts Due to Underlying Changes in the Fed Cattle Industry." *Extension Farm and Ranch Management News.* doi: 10.13014/frm00014.

Heckman, J. J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47(1):153–161. doi: 10.2307/1912352.

Heckman, J. J., and B. E. Honoré. 1990. "The Empirical Content of the Roy Model." *Econometrica* 58(5):1121–1149. doi: 10.2307/2938303.

Henderson, G. 2022, March 28. "Senators Revise Cattle Price Discovery and Transparency Act." *AgWeb.* Available online at https://www.agweb.com/news/livestock/beef/senators-revise-cattle-price-discovery-and-transparency-act.

Ji, I., C. Chung, and J. Lee. 2017. "Measuring Oligopsony Power in the U.S. Cattle Procurement Market: Packer Concentration, Cattle Cycle, and Seasonality." *Agribusiness* 33(1):16–29. doi: 10.1002/agr.21490.

Koontz, S. R. 2010. *What Does the RTI Study Say about Captive Supplies in the Cattle and Beef Industry?* Agricultural Marketing Report 2. Fort Collins, CO: Colorado State University Department of Agricultural and Resource Economics. Available online at https://mountainscholar.org/handle/10217/44560.

———. 2015. "Marketing Method Use in Trade of Fed Cattle: Causes and Consequences of Thinning Cash Markets and Potential Solutions." Paper presented at the joint annual meeting of the American Applied Economics Association & Western Agricultural Economics Association, July 26–28, San Francisco, California.

Lee, L.-F. 1995. "The Computation of Opportunity Costs in Polychotomous Choice Models with Selectivity." *Review of Economics and Statistics* 77(3):423–435. doi: 10.2307/2109905.

Mallory, S., E. A. DeVuyst, K. C. Raper, D. Peel, and G. Mourer. 2016. "Effect of Location Variables on Feeder Calf Basis at Oklahoma Auctions." *Journal of Agricultural and Resource Economics* 41(3):1–13. doi: 10.22004/ag.econ.246171.

McDonald, R. A., and T. C. Schroeder. 2003. "Fed Cattle Profit Determinants under Grid Pricing." *Journal of Agricultural and Applied Economics* 35(1):97–106. doi: 10.1017/S1074070800005964.

McGuire, D. P. 1986. "The Occurrence of an Increase in Correlation with Explicit Selection." *Psychometrika* 51(2):331–333. doi: 10.1007/BF02293989.

Peel, D. S., D. Anderson, J. Anderson, C. Bastian, S. Brown, S. R. Koontz, and J. Maples. 2020. *Fed Cattle Price Discovery Issues and Considerations.* Fact Sheet E-1053. Stillwater, OK: Oklahoma State University. Division of Agricultural Sciences and Natural Resources, Cooperative Extension Service. Available online at https://extension.okstate.edu/fact-sheets/print-publications/e/fed-cattle-price-discovery-issues-and-considerations-e-1053.pdf.

Peel, M. J., and G. H. Makepeace. 2012. "Differential Audit Quality, Propensity Score Matching and Rosenbaum Bounds for Confounding Variables." *Journal of Business Finance & Accounting* 39(5-6):606–648. doi: 10.1111/j.1468-5957.2012.02287.x.

Roy, A. D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3(2):135–146. doi: 10.1093/oxfordjournals.oep.a041827.

Sabasi, D. M., C. T. Bastian, D. J. Menkhaus, and O. R. Phillips. 2013. "Committed Procurement in Privately Negotiated Markets: Evidence from Laboratory Markets." *American Journal of Agricultural Economics* 95(5):1122–1135. doi: 10.1093/ajae/aat051.

SAS Institute, Inc. 2014. *SAS/ETS® User's Guide, Version 13.2*. Cary, NC: SAS Institute, Inc. Available online at https://support.sas.com/documentation/cdl/en/etsug/67525/HTML/default/viewer.htm#titlepage.htm.

Schroeter, J. R., and A. Azzam. 2003. "Captive Supplies and the Spot Market Price of Fed Cattle: The Plant-Level Relationship." *Agribusiness* 19(4):489–504. doi: 10.1002/agr.10070.

US Department of Agriculture. 2021. "Prices Received for Cattle by Month – United States." Washington, DC: USDA National Agricultural Statistics Service. Available online at https://www.nass.usda.gov/Charts_and_Maps/Agricultural_Prices/priceca.php [Accessed May 1, 2021].

———. 2022. "Sector at a Glance." Washington, DC: USDA Economic Research Service. Available online at https://www.ers.usda.gov/topics/animal-products/cattle-beef/sector-at-a-glance/ [Accessed May 28, 2022].

Ward, C. E. 1999. *Packer Concentration, Captive Supplies, and Their Impacts: A Review*. Stillwater, OK: Oklahoma State University. Division of Agricultural Sciences and Natural Resources Cooperative Extension Service.

Ward, C. E., S. R. Koontz, and T. C. Schroeder. 1998. "Impacts from Captive Supplies on Fed Cattle Transaction Prices." *Journal of Agricultural and Resource Economics* 23(2):494–514. doi: 10.22004/ag.econ.31205.

Ward, C. E., T. C. Schroeder, and D. M. Feuz. 2017. *Grid Pricing of Fed Cattle: Base Prices and Premiums-Discounts*. AGEC 560. Stillwater, OK: Oklahoma Cooperative Extension Service. Available online at http://pods.dasnr.okstate.edu/docushare/dsweb/Get/Document-1975/AGEC-560web.pdf.

Wimmer, B. S., and B. Chezum. 2003. "An Empirical Examination of Quality Certification in a "Lemons Market"." *Economic Inquiry* 41(2):279–291. doi: 10.1093/ei/cbg007.

———. 2006. "Adverse Selection, Seller Effort, and Selection Bias." *Southern Economic Journal* 73(1):201–218. Publisher: Southern Economic Association. doi: 10.2307/20111883.

Winship, C., and R. D. Mare. 1992. "Models for Sample Selection Bias." *Annual Review of Sociology* 18(1):327–350. doi: 10.1146/annurev.so.18.080192.001551.

Xia, T., J. M. Crespi, and K. C. Dhuyvetter. 2019. "Could Packers Manipulate Spot Markets by Tying Contracts to Future Prices? And Do They?" *Canadian Journal of Agricultural Economics* 67(1):85–102. doi: 10.1111/cjag.12179.

Xia, T., and R. J. Sexton. 2004. "The Competitive Implications of Top-of-the-Market and Related Contract-Pricing Clauses." *American Journal of Agricultural Economics* 86(1):124–138. doi: 10.1111/j.0092-5853.2004.00567.x.

Zhang, T., and B. W. Brorsen. 2010. "The Long-Run and Short-Run Impact of Captive Supplies on the Spot Market Price: An Agent-Based Artificial Market." *American Journal of Agricultural Economics* 92(4):1181–1194. doi: 10.1093/ajae/aaq033.