Uncovering Factors Associated with Price Ranges from Fed Cattle Negotiated Cash Sales using Machine Learning

Zuyi Wang[1], Hernan Tejeda[2], and Man-Keun Kim[3]

[1] Department of Applied Economics, Utah State University, email: zuyi.wang@usu.edu

[2] Department of Agricultural Economics and Rural Sociology, University of Idaho, email: htejeda@uidaho.edu

[3] Department of Applied Economics, Utah State University, email: mk.kim@usu.edu

Selected Poster prepared for presentation at the 2023 Agricultural & Applied Economics Association Annual Meeting, Washington DC: July 23- 25, 2023

Uncovering Factors Associated with Price Ranges from Fed Cattle Negotiated Cash Sales using Machine Learning

Zuyi Wang[1], Hernan A. Tejeda[2], and Man-Keun Kim[3]

1. Motivation

The results of policy proposals to raise negotiated cash sales in the U.S. cattle market, e.g., "Spot Market" bill and "Cattle Market Transparency Act" of 2021, are uncertain.

Recent wide spreads between boxed beef cutout values and fed cattle prices have reignited beef cattle producers' long-held concerns about price discovery (Boyer et al., 2022).

This study is complementary work to Boyer et al. (2022) using Machine Learning (ML) to learn hidden information and/or patterns behind the fed cattle negotiated cash price data with volume, day of the week, sex, grade, weight range, and other factors.

Why ML? Regression methods may suffer from the followings (Storm et al., 2020)

- Restrictive functional forms with little theoretical ground,
- Limited ability to extract information from unstructured data,
- Limited ability to deal with a large number of explanatory variables, and
- Causal inference and identification problem

2. Data

Daily negotiated cash purchase data of fed cattle in Texas/Oklahoma/New Mexico, Kansas, Nebraska, and Iowa/Minnesota from 2010 to 2019.

Price range is defined as the difference between daily high price and daily low price (Boyer et al., 2022); average price range is $1.15/cwt (Fig 1).

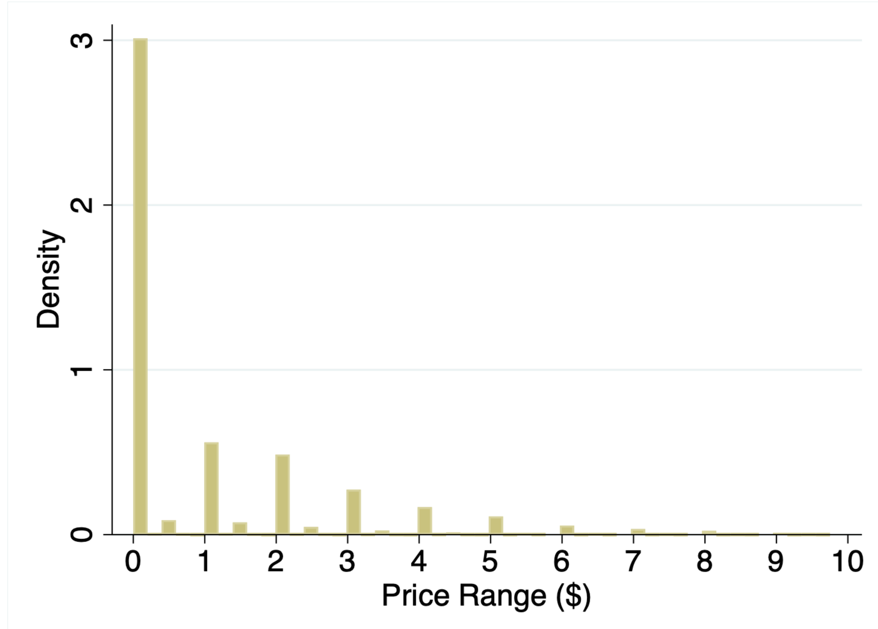

Fig 1. Price Range

Predictors and variables of interest (see Fig 2. frequency table of characteristics for more information)

pd.eq.0: price range is zero
LiveFOB: Live free on board
dress_del: dressed delivered
Head ≤ 76: trade volume ≤ 76 cows
Kansas: trade at Kansas
wgt>1250: average weight greater than 1250 lb.

pd.gt.2: price range is greater than $2/cwt
Head ≥ 684: trade volume ≥ 684,
choice ≥ 80%: beef grading 80% or higher,
800≤wgt≤875: average weight between 800,
IA_MN: trade at Iowa/Minnesota

Not shown in Fig 2 and Fig 3, date of the week, month, type of cattle are included.

3. Association Rule Method (ARM) and Interpretations

ARM finds "lift" to measures how much more likely Y show up given X (Agrawal et al., 1993).

$$\text{Lift} = \frac{\text{condfidence}(\text{X} \rightarrow \text{Y})}{\text{support}(\text{Y})}$$

where $\text{support} = \frac{\text{count}(\text{X})}{\text{N}}$, that is, frequency of X and

$\text{condfidence}(\text{X} \rightarrow \text{Y}) = \frac{\text{support}(\text{X},\ \text{Y})}{\text{support}(\text{X})}$, which measures likelihood that Y on right hand side shows up with X on left hand side (see Fig 3 for value of support and lift).

A large lift value is a strong indicator that a rule (set of variables to connect to right hand side variable, price range, for example, `{heifer, LiveFOB, head ≤ 76} → {pd.eq.0}` (Fig. 4)

lift > 1 indicates that the two variables (characteristics) in the rule occur more often together than what would be expected if they were independent of each other. For example, lift of `{choice.over.80%, steer} →{pd.gt.2}` in Fig. 5 is 3.85, which means that transaction with choice over 80% cattle and steer would be **3.85 times more likely** to have price range larger than $2/cwt.

4. Result

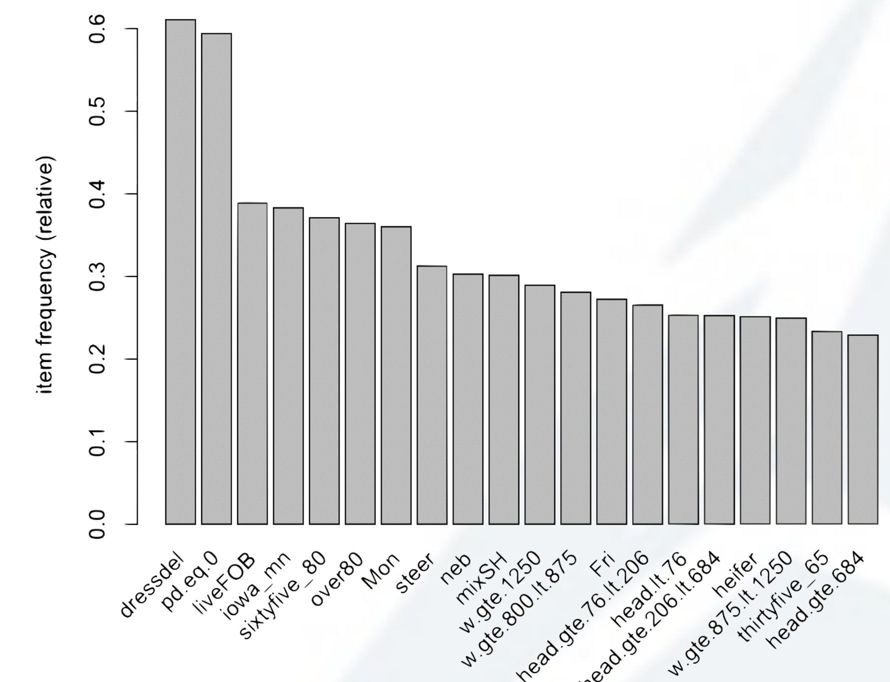

Fig 2. Characteristics frequency table

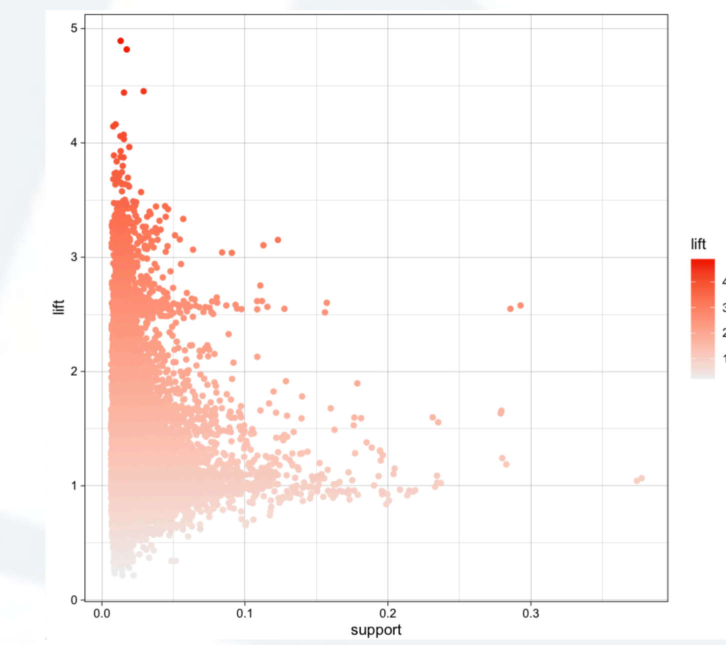

Fig 3.Scatter plot for all rules

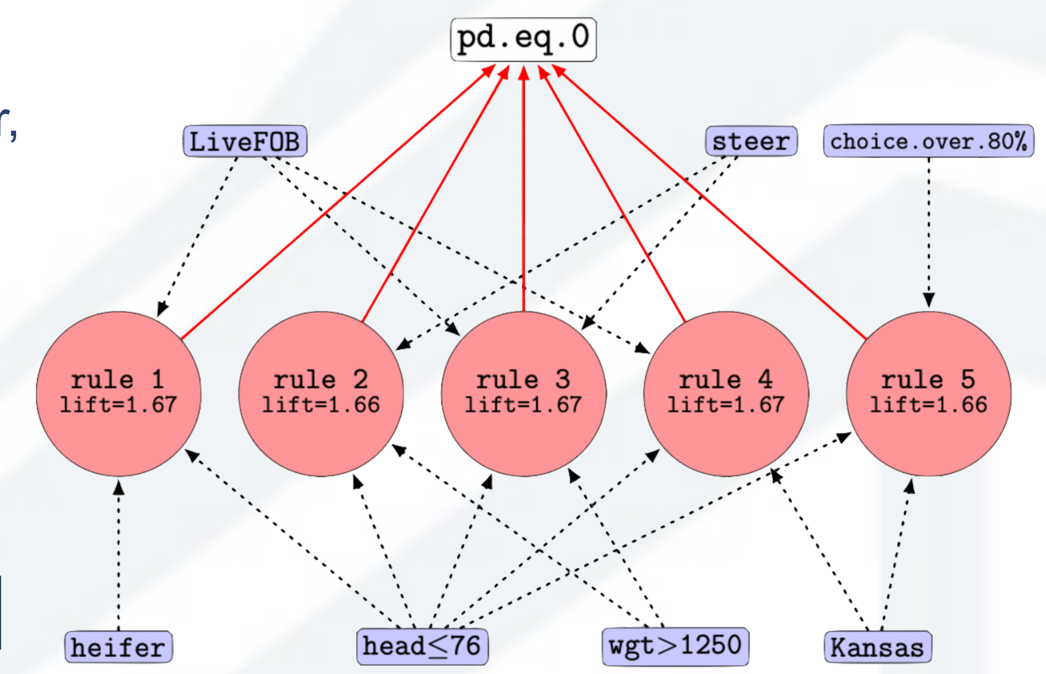

Fig 4. Top 5 rules to pd.eq.0

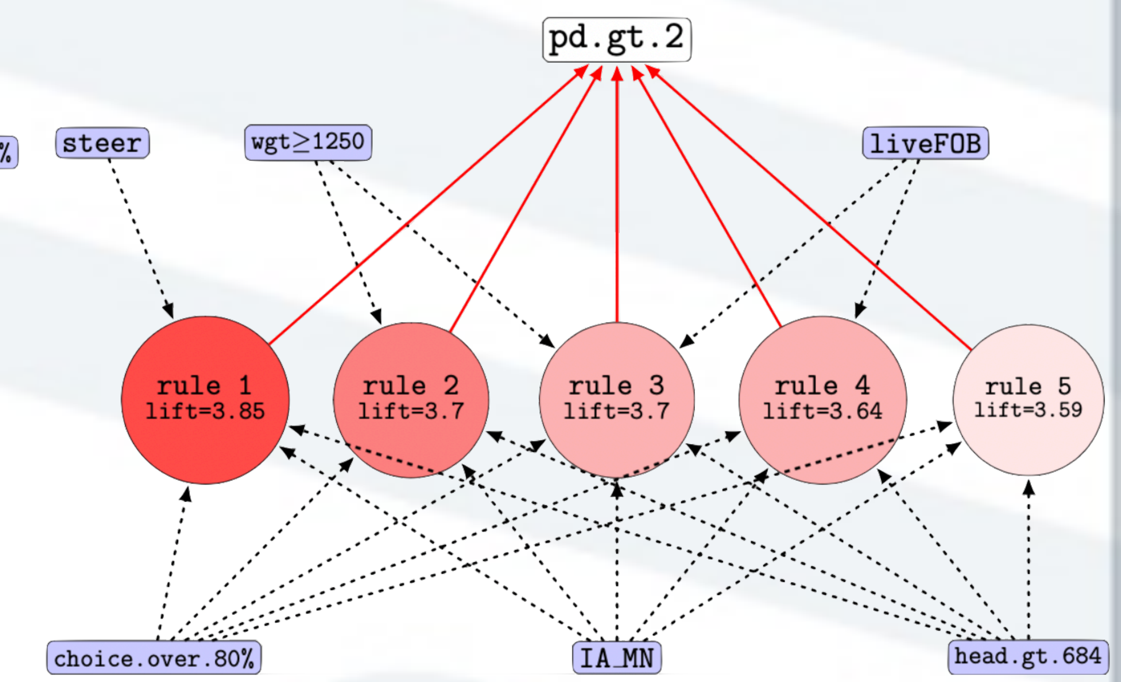

Fig 5. Top 5 rules to pd.gt.2

5. Discussion

Note that values of lift for each rule (1 through 5) are presented in Figs 4 and 5.

For price ranges being zero, in Fig 4, the lift for Rule 1 is 1.67, which implies {heifer, liveFOB, head ≤ 76} would lead to being 1.67 times more likely to have zero price difference than other transaction conditions. Similar results for Rule 3 from {head ≤ 76, wgt>1250, Live FOB, steer} are obtained.

In Figure 4, when the price ranges are zero, Rule 1 shows a lift of 1.67. This means that the condition {heifer, liveFOB, head ≤ 76} is 1.67 times more likely to result in zero price difference compared to other transaction conditions. Similar results were observed for Rule 3, which includes {head ≤ 76, wgt > 1250, Live FOB, steer}

This study complements Boyer et al. (2022) by determining which factors (variables of interest) may have a larger effect on the dispersion level of daily fed cattle transaction data. It serves to provide guidance on which fed cattle transaction characteristics may be more suitable to identify markets with more (or less) daily price variability.

Here we identify variables of interest from fed cattle transaction data, that may lead to increased chances of obtaining "targeted" price ranges (e.g. price ranges being zero, prices ranges being above $2 / cwt, etc.), enabling market analysts as well as policy makers to converge on particular characteristics for their analysis efforts.

Authors

1. **Zuyi Wang**, Research Assistant, Applied Economics, Utah State University
2. **Hernan A. Tejeda**, Associate Professor, Agricultural Economics and Rural Sociology, University of Idaho
3. **M-K. Kim**, Professor, Applied Economics, Utah State University