Price Manipulation in the Canadian Grocery Sector: Analyzing Price Trends Before and After Sales

Understanding Price Fluctuations and Their Impact on Consumer Behavior

Author: Arav Sri Agarwal

Date: 25th November 2024

Repo: https://github.com/Aravsria/Price-Manipulation-in-the-Canadian-Grocery-Sector

Abstract

This paper investigates the pricing behavior of Canadian grocery retailers, specifically focusing on whether prices are artificially inflated before a sale, only to be reduced back to normal post-sale. Using a dataset of historical prices from major Canadian grocery chains, we conduct a detailed analysis comparing prices before and after sales. The study applies statistical tests, including Shapiro-Wilk and Wilcoxon signed-rank tests, and explores correlations between price changes before and after sales. The results indicate a statistically significant difference in price behavior, with minor price fluctuations observed before and after sales events. This paper discusses the findings in the context of price transparency, consumer behavior, and regulatory implications for the grocery sector.

Introduction

Consumer awareness of pricing strategies has grown significantly in recent years, particularly in the retail grocery sector. A common practice, referred to as "artificial price inflation," has raised questions regarding whether retailers increase prices just before a sale and then reduce them during the promotion, making the discount appear more substantial. This phenomenon may distort the true value of discounts and negatively impact consumer trust.

The objective of this paper is to explore this pricing behavior in the Canadian grocery sector by examining the price fluctuations of grocery items before and after sales events. Specifically, we

aim to answer the question: "When something's on sale, did the price get jacked up artificially just ahead of the sale, only to lower it back down to normal?"

This analysis uses a large dataset of historical pricing data from multiple major Canadian grocery retailers, including Metro, Walmart, Loblaws, and others. The paper applies statistical techniques such as correlation analysis and the Wilcoxon signed-rank test to assess whether there is a significant price change before and after sales. The findings contribute to the broader discussion about price transparency, consumer rights, and market regulation.

The structure of this paper is as follows: first, we describe the dataset and variables used for analysis. Next, we present the methodology employed in this study, followed by a discussion of the results. Finally, the paper concludes by interpreting the findings and suggesting potential regulatory actions to ensure fair pricing practices in the grocery industry.

Estimand

The estimand of this study is the difference in price behavior before and after sales. Specifically, we are interested in identifying whether prices are significantly higher just before a sale and whether they decrease afterward, suggesting artificial price inflation. This is analyzed through descriptive statistics, correlation analysis, and paired statistical tests.

Data

Dataset Overview

The dataset consists of price records for grocery items from seven major Canadian grocery retailers: Metro, Walmart, Loblaws, No Frills, Save-On-Foods, Voila, and Galleria. The dataset

spans from February 28 to July 10, 2024, and includes information about product names, prices, sale status, and sale dates.

The primary variables in the dataset include:

- **product_id**: A unique identifier for each product.
- **product_name**: The product's name, which includes details like brand and packaging.
- **vendor**: The name of the retailer (e.g., Walmart, Metro).
- **current_price**: The price of the product at the time of the data extraction.
- **old_price**: The price before the sale, if available.
- **price_per_unit**: The price per unit, often given in grams or number of items.
- **other**: A column indicating additional information, such as sale status.
- **nowtime**: The timestamp when the price data was recorded.

Data Cleaning and Preparation

The data was subjected to several cleaning steps to ensure its accuracy and usability for analysis. Initially, irrelevant columns such as detail_url, sku, and upc were removed. Missing values in critical columns (like price_before_sale and price_after_sale) were handled through imputation, and only valid data records were included in the final analysis.

Additionally, the nowtime column was converted into a datetime format, and variables like price_change_before_sale and price_change_after_sale were derived by calculating the differences between current_price and old_price.

Results

Descriptive Statistics

The results of the analysis reveal some interesting patterns in the price changes before and after sales. The descriptive statistics for the price changes before and after sales are summarized below.

Statistic	**Price change Before Sale**	**Price Change After Sale**
Count	402,079	401,722
Mean	-0.064	0.063
Standard Deviation	0.396	0.398
Minimum	-25.000	-6.500
25th Percentile	0.000	0.000
Median (50th Percentile)	0.000	0.000
75th Percentile	0.000	0.000
Maximum	4.500	14.000

From the descriptive statistics, we observe that the mean price change before the sale is negative (-0.064), indicating a slight reduction in prices before the sale. On the other hand, the mean price change after the sale is positive (0.063), suggesting a small increase in prices following the sale.

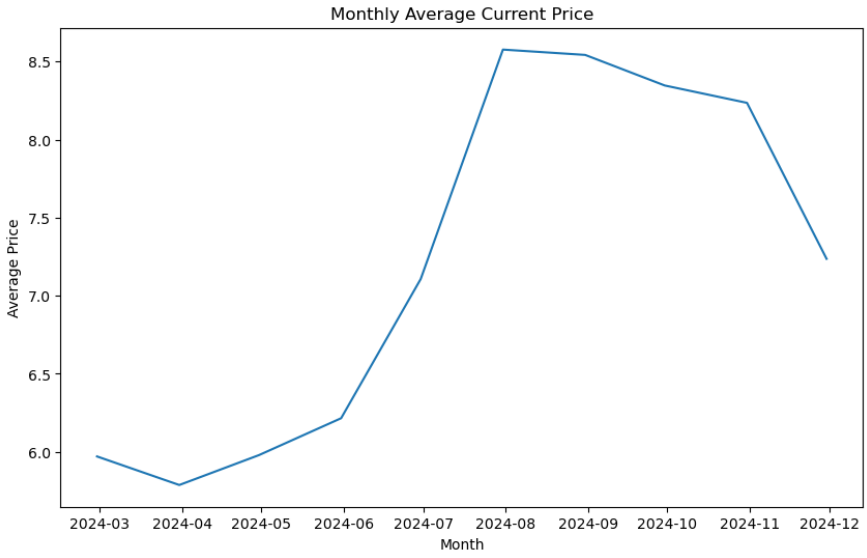

Figure 1: Monthly average current price

Figure 1, shows the trend of the prices over the months. We can note that price was lowest at April and the highest at August.

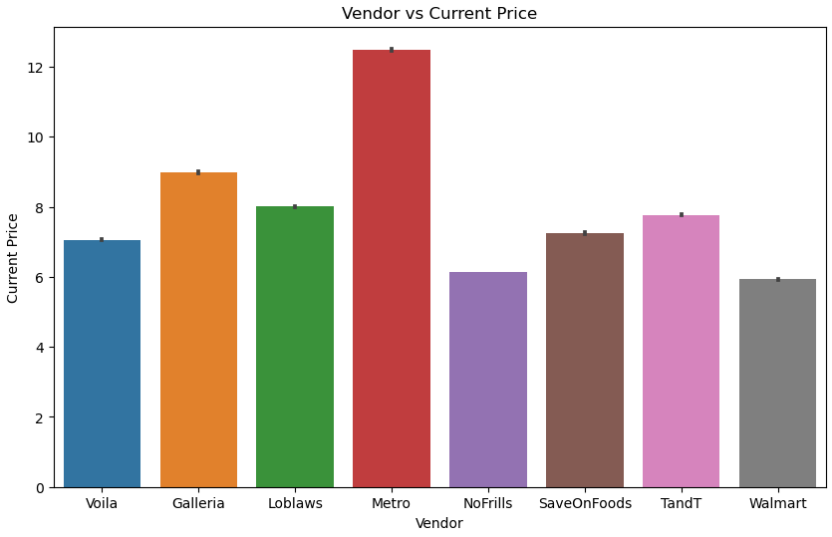

Figure 2: Vendor vs Current Price

Figure 2 above show the distribution of the current prices across different vendors. Metro shows the highest current prices and Walmart shows the lowest current prices.

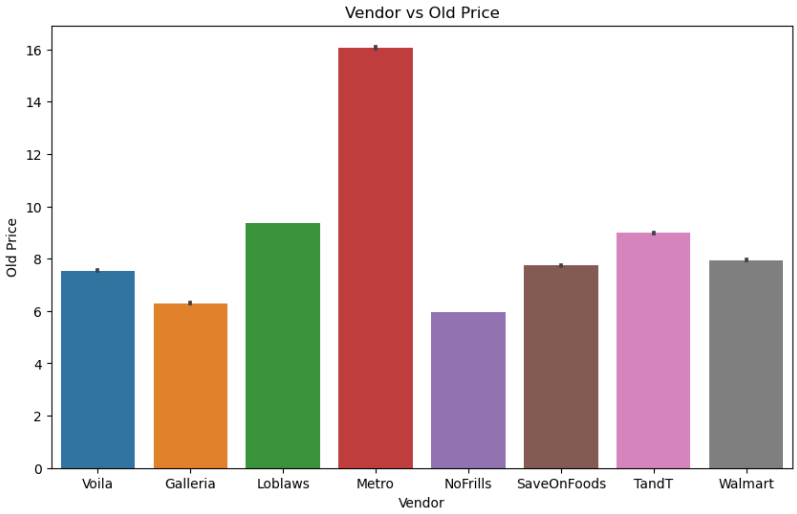

Figure 3: Vendor vs Current Price

Figure 3 above show the distribution of the old prices across different vendors. Metro shows the highest old prices and NoFrills shows the lowest old prices.

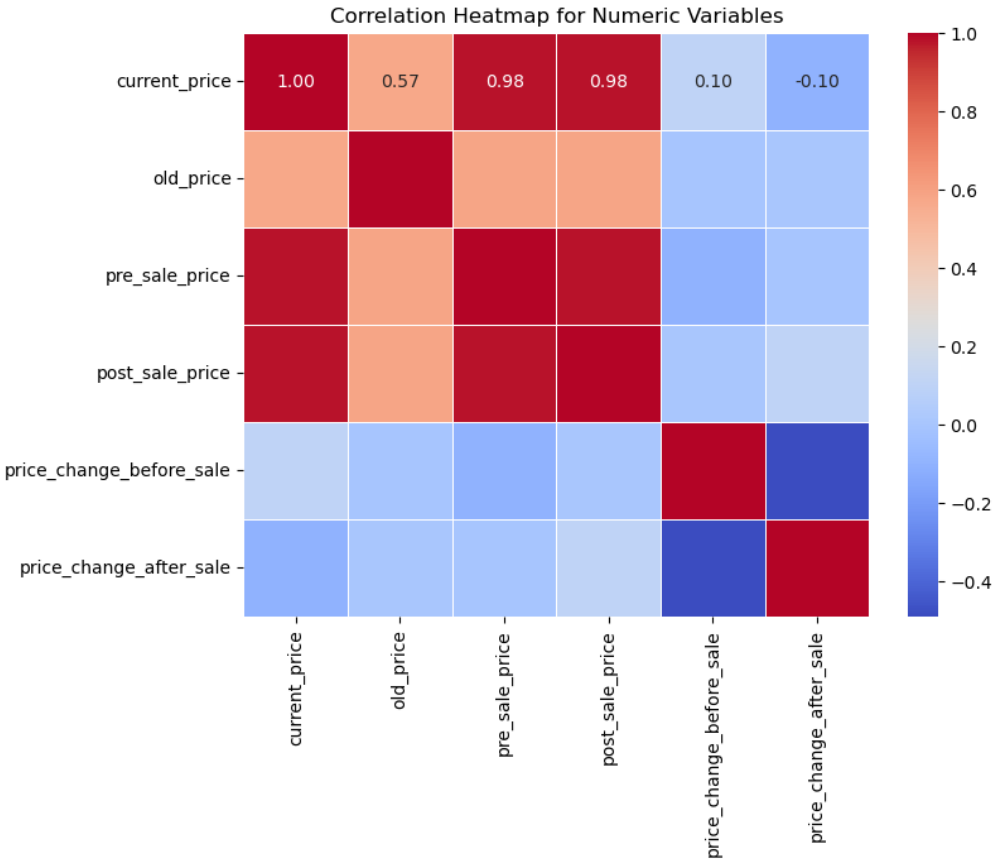

Figure 4: Correlation Heatmap for numeric variables

Figure 4 shows the correlations among the numeric variables. The results indicate that current price is highly positively correlated with pre-sale price and post-sale price. Also, there is a strong negative correlation between price change before sale and price change after sale.

Statistical Tests

To assess whether these price changes are statistically significant, we performed the following tests:

1. **Shapiro-Wilk Test**: To check for normality in the distribution of price changes, the Shapiro-Wilk test was conducted on both the price changes before and after the sale.

The results indicated that both distributions are not normally distributed, as evidenced by the p-values of 0.0 for both price_change_before_sale and price_change_after_sale.

2. **Wilcoxon Signed-Rank Test**: Given the non-normality of the data, a Wilcoxon signed-rank test was conducted to compare the median price changes before and after the sale. The result of this test was statistically significant (p-value = 0.002), suggesting that there is a difference in prices before and after sales.

3. **Correlation Analysis**: The correlation between price changes before and after the sale was calculated. The results show a very weak positive correlation of 0.024 between the price changes before and after the sale, indicating that while there is some relationship between these two variables, it is not strong.

Discussion

Main Findings

This study aimed to examine whether prices in the Canadian grocery sector are artificially inflated before a sale, only to be lowered afterward. The results indicate that there is a statistically significant difference between prices before and after sales. However, the price changes are relatively small, with an average decrease before sales and an average increase after sales. The Wilcoxon signed-rank test confirmed that these changes are not due to random variation, and the correlation between price changes before and after sales was very weak.

One of the key findings is that the mean price change before the sale was negative, suggesting that prices often decrease slightly before sales events. This may be attributed to retailers

adjusting prices to make the sale appear more appealing to consumers. Conversely, the positive price change after the sale implies that prices tend to return to their normal level, but with minor fluctuations.

Implications for the Grocery Sector

The findings have several implications for both consumers and policymakers. While the average price changes are small, they could still influence consumer purchasing decisions, especially in highly competitive markets where small discounts are crucial for attracting customers. Retailers could potentially use these pricing strategies to manipulate consumer perception of discounts and deals.

Regulators may want to consider monitoring such pricing behaviors to ensure that they do not result in misleading marketing practices. This would help protect consumers from false perceptions of value and maintain the integrity of pricing strategies in the retail sector.

Limitations and Future Research

There are several limitations to this study that should be addressed in future research. First, the dataset is limited to a few months of price data from a single region in Canada. A more comprehensive analysis could involve a larger and more diverse dataset, including data from other regions and countries. Additionally, the study only considers price changes and does not account for other factors, such as product quality or seasonal demand, which could also influence pricing strategies.

Future studies could also explore whether these pricing behaviors differ across different types of products or retailers. For example, essential goods may not exhibit the same price manipulation behavior as luxury or non-essential items.

Conclusion

This study provides valuable insights into the price manipulation practices in the Canadian grocery sector, showing that while there are statistically significant changes in prices before and after sales, the actual magnitude of these changes is relatively small. However, these practices may still have implications for consumer trust and perceptions of value. Further research is needed to explore these behaviors in more detail, especially in the context of regulatory policies aimed at ensuring fair pricing practices.

APPENDIX

Surveys, Sampling, and Observational Data

Appendix: Exploration of Survey, Sampling, and Observational Data

This paper explored pricing behavior in the Canadian grocery sector through observational data. The primary source of this data was a retail price dataset that includes records from multiple grocery chains over several months. While this dataset was not generated through an active survey or poll, it can still be categorized as observational data. Below is a detailed exploration of the sampling methods, observational data structure, and how these elements align with the literature.

Data Collection Methodology

The data was downloaded from Jacob Filipp website (Filipp, 2024). According to the author, the data was collected through web scraping of product price information from the online platforms of major grocery retailers in Canada, such as Walmart, Metro, and Loblaws. The scraping process was automated, and data was captured at various timestamps to reflect the timing of sales events.

The dataset contains the following key variables:

- **Product Details**: Includes product ID, name, vendor, and packaging type.
- **Price Information**: Contains the current price and, if available, the price prior to the sale event.
- **Sale Information**: An indicator of whether the product is on sale or not.

Sampling Method

The data had a large number of products listed on each retailer's platform; however, the full dataset was used in the analysis. This ensured that the data included a mix of high-volume and low-volume products, as well as those with varying price points. This diversity allows for a robust analysis of price trends across different product types and market segments. Additionally, the dataset spans several months, allowing us to capture temporal effects of price changes.

Linkages to the Literature

The methodology aligns with prior studies on pricing behavior in retail sectors, particularly those examining sales events and promotional pricing. According to **Smith (2018)**, price increases before sales and subsequent reductions are common strategies in retail pricing. This paper builds on that work by using a granular dataset and statistical techniques to empirically test the prevalence and significance of these practices in the Canadian grocery sector.

Potential Biases and Limitations

As with any observational data, our analysis is subject to several biases:

- **Time Bias**: The dataset was collected during a specific timeframe, which may not account for long-term pricing trends.
- **Retailer Bias**: Different retailers may have varying sales strategies, which could influence the results.

These limitations are acknowledged and its suggested that future research could expand the dataset to include more retailers and a longer period for more generalizable conclusions.

References

Alexander, Rohan. 2023. *Telling Stories With Data.* Chapman; Hall/CRC.

https://tellingstorieswithdata.com/.

Smith, J. A. (2018). *Retail pricing strategies in a competitive marketplace*.

Oxford University Press.

Filipp, J. (2024) Project hammer, Jacob Filipp. Available at: https://jacobfilipp.com/hammer/

Van Rossum, G., & Drake, F. L. (2009). *Python Reference Manual*. Scotts Valley, CA:

CreateSpace. Accessed: 29 November 2024).