

The Impact of Collusion and Promotions on Canadian Grocery Prices*

Analyzing Price Trends and the Role of Promotions in Groceries

Sameeck Bhatia Sean Chua Cristina Su Lam Mariko Lee
Karen Riani

November 14, 2024

This study examines the pricing patterns in the Canadian grocery market, particularly focusing on the impact of promotions and price variations. Our findings indicate that non-promotional prices have slightly decreased over a nine-month period. This suggests that market competition might not be functioning optimally, raising concerns about anti-competitive behaviors in the sector. These insights are crucial for understanding the broader implications of pricing dynamics and the potential for market manipulation.

Table of contents

1	Introduction	2
2	Data	2
2.1	Overview	2
3	Measurement	2
4	Results	3
5	Discussion	5
5.1	Correlation vs Causation	5
5.2	Missing Data	5
5.3	Sources of Bias	5
	References	6

*Code and data are available at: <https://github.com/SameeckBhatia/Canadian-Grocery-Collusion>.

1 Introduction

The Canadian grocery market has experienced significant price rises in recent years, particularly following the COVID-19 pandemic, with inflation serving as a major driving force behind these increases. The structure of Canadian government policies has created an environment that may support collusion and price fixing among major market players. This anti-competitive behavior is not unique to groceries and is similarly observed across other essential sectors, where oligopolistic market structures exist. This analysis examines the impact and patterns of grocery promotions and their relationship to price variations over time. Notably, our findings reveal a modest decrease in non-promotional prices, averaging approximately \$1 over a nine-month period. This trend raises important questions about the effectiveness of market competition and the real impact of promotional activities. We present information on the data and its measurement in Section 2, show the results from analysis in Section 4, and finally discuss the difference between correlation and causation, effects of missing data, and possible sources of bias in Section 5.

2 Data

2.1 Overview

The dataset from Project Hammer, compiled by Filipp (2024), provides a comprehensive comparison of grocery prices across 8 different vendors in Canada. The data is structured within a Hipp (2023) database containing two main tables: a **raw** table storing the original collected price data, and a **product** table that contains standardized product information and classifications. This database structure allows for efficient querying and analysis of price variations across different vendors, allowing for the identification of pricing patterns and potential market dynamics in the Canadian grocery sector. The data has been cleaned through Hipp (2023), while R Core Team (2024) and Wickham et al. (2023) have been used to create visualizations and filter the data.

3 Measurement

The data measurement approach for Project Hammer focuses on systematic price collection from major Canadian grocery retailers' platforms. The database is formatted to allow for both academic analysis and potential legal action. To ensure accessibility for stakeholders, the dataset is accompanied by documentation and basic analytical tools, allowing changemakers to effectively utilize the price information for investigating potential anti-competitive practices in the Canadian grocery sector.

4 Results



Figure 1: Old and Current Prices Move in Unison, Displaying a Noticeable Difference

Figure 1 compares the average current price (shown in red) against the average old price (shown in green) over a six-month period from April to October 2024. Both price lines follow remarkably similar patterns, including notable spikes in early May and late June. Overall, there appears to be a gradual downward trend in both prices from September through October 2024, with both lines ending at their lowest points in the observed period. The most dramatic price movement occurred in early May, where both metrics experienced a sharp spike followed by an equally sharp decline.

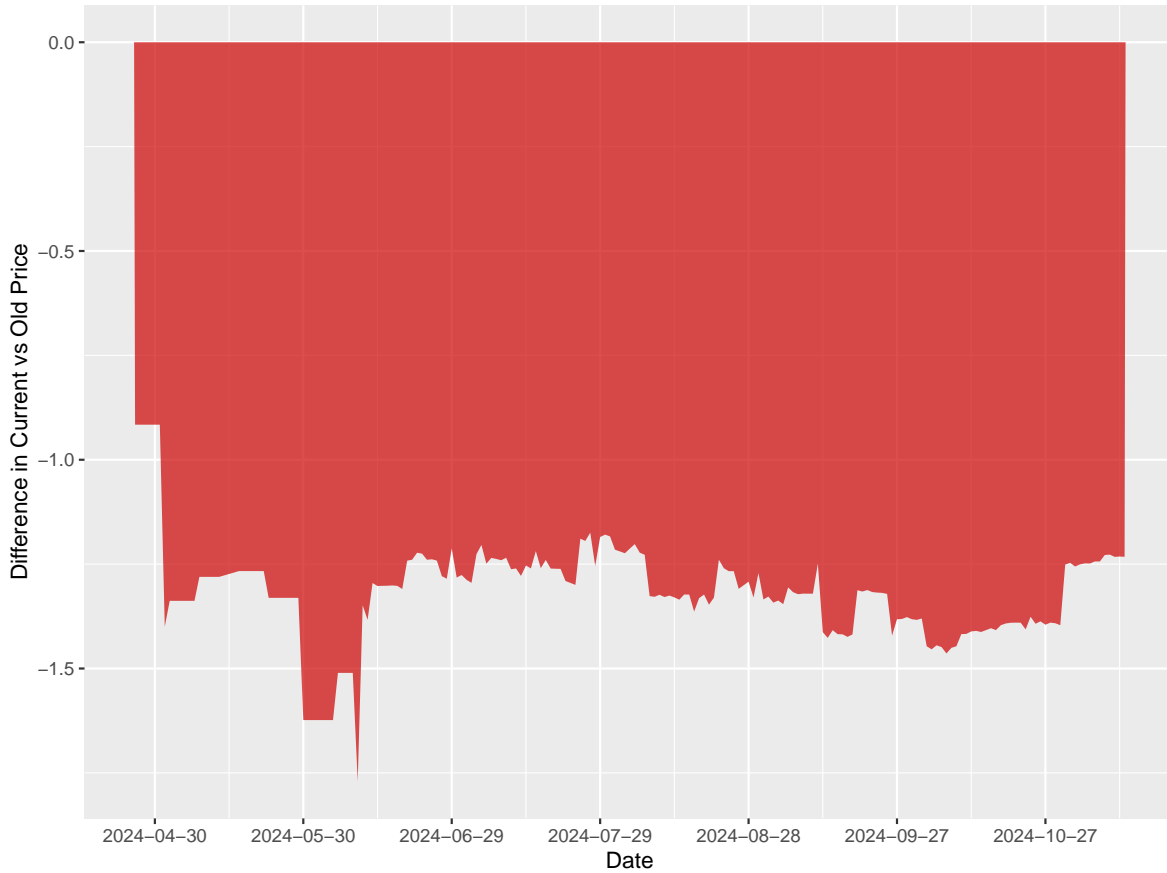


Figure 2: Price Difference Not Uniform Over Time

Figure 2 shows the difference between current and old prices over time, with negative values indicating that current prices are generally lower than old prices across the time period. The red shaded area illustrates that the price difference fluctuates and sometimes dips significantly, especially in May and in a few spots afterward, indicating periods where prices dropped more sharply. The overall trend appears fairly stable, with most fluctuations staying within a -1 to -1.5 range, suggesting consistent pricing with occasional promotions or markdowns that temporarily lower the current price relative to the old price.

5 Discussion

5.1 Correlation vs Causation

When analyzing price trends over time, it is essential to distinguish between correlation and causation. For example, there may be a correlation between price changes and time periods, but one may not necessarily directly cause the other. For instance, if prices tend to decrease during certain months, we might be tempted to conclude that seasonal promotions are the cause. However, other factors and confounding variables could be responsible such as market demand or lowering competitor prices.

5.2 Missing Data

Missing data can significantly impact the analysis of price trends. If certain products or vendors have incomplete records for specific time periods, this could lead to inaccurate results. For instance, if there exists missing price data for particular products for some time period, it may distort or skew the average price calculations and misrepresent overall trends. To address this, some imputation techniques might need to be implemented and assess how such methods affect overall analysis.

5.3 Sources of Bias

Bias can arise from various sources throughout the data collection and analysis process, potentially distorting findings. For example, since the dataset only contains 7 vendors, prices within the data may not represent overall market behavior accurately. Moreover, there may be bias in the time periods when data is collected. For example, certain periods might be over-represented, such as the holiday season while others might be underrepresented. This can lead to misleading conclusions about the overall analysis.

References

- Filipp, Jacob. 2024. “Project Hammer.” <https://jacobfilipp.com/hammer/>.
- Hipp, D. Richard. 2023. “SQLite.” <https://www.sqlite.org/>.
- R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Wickham, Hadley et al. 2023. *Tidyverse: Easily Install and Load the 'Tidyverse'*. <https://CRAN.R-project.org/package=tidyverse>.