Cross-Vendor and Temporal Pricing Trends in Beverages*

A Bayesian Analysis of Canadian Grocery Data

Betty (Zaihua) Liu

December 6, 2024

This paper investigates pricing trends in the Canadian grocery sector, focusing on beverages across six major vendors. Using Bayesian modeling, we identify vendor-specific strategies, seasonal effects, and the influence of historical prices on current pricing. The analysis reveals significant differences between budget-friendly vendors like Walmart and premium retailers like Galleria, as well as seasonal markdown patterns in the fall months. These findings provide actionable insights for consumers, vendors, and policymakers, promoting competition and transparency in the grocery market.

1 Introduction

The Canadian grocery sector is a dynamic and competitive market, where pricing strategies play a critical role in shaping consumer behavior and influencing market outcomes. Vendors frequently adjust their pricing to respond to regional differences, seasonal demand, and competitive pressures, creating notable disparities in prices for similar products. Understanding these pricing behaviors is essential for promoting transparency, fairness, and competition in the market.

This paper focuses on analyzing beverage pricing trends, a product category that reflects diverse pricing strategies and consumer demand patterns. By examining vendor-specific and temporal effects on current prices, we aim to uncover insights into how prices are set, adjusted, and influenced over time. Leveraging Bayesian modeling, this research estimates the influence of historical prices, vendor characteristics, and monthly variations on current prices, offering a probabilistic framework that accounts for uncertainty in the data.

*Code and data are available at: https://github.com/betty-tGM2/Analysis-of-Canadian-Grocery-Store-Beverage-Prices.git.

The study is motivated by the broader objectives of Project Hammer, a national initiative aimed at reducing collusion and fostering competition in the Canadian grocery sector. By shedding light on vendor-specific strategies and seasonal pricing behaviors, this research addresses gaps in understanding vendor dynamics, particularly in a market where pricing disparities can hinder consumer welfare and market efficiency. Beverages are chosen as a representative category for their ubiquity and their pricing diversity across vendors and time.

Key findings reveal significant variations in pricing strategies among vendors, with Walmart consistently maintaining lower prices, while Galleria exhibits the highest variability. Temporal trends suggest seasonal pricing adjustments, particularly from June to November, as vendors respond to demand shifts. The Bayesian model used demonstrates strong predictive performance, capturing vendor-specific and temporal dynamics effectively.

This research contributes to the growing discourse on market competition by providing actionable insights for consumers, vendors, and policymakers. For consumers, understanding pricing trends enhances decision-making; for vendors, it provides a benchmark for competitive positioning; and for policymakers, it offers a foundation for developing regulations that promote market fairness and competition.

The remainder of this paper is structured as follows: Section Section 2 introduces the dataset, including simulation and preprocessing steps; Section **?@sec-methods** details the exploratory data analysis and Bayesian modeling framework; Section **?@sec-results** presents the findings, focusing on vendor-specific and temporal pricing trends; and Section **?@sec-discussion** interprets these results, their implications, and the study's limitations. Finally, Section **?@sec-conclusion** summarizes the contributions and suggests avenues for future research.

2 Data

2.1 Overview

This study uses simulated and cleaned data using statistical programming language R (R Core Team 2023) to do the data analysis and visualizations. Following the guidelines of Alexander (2023) (Alexander 2023), the dataset was rigorously cleaned and transformed to ensure high-quality analysis. The cleaning process removed invalid prices, ensured consistency in formatting, and excluded non-relevant products. Simulated data techniques were employed to fill gaps and emulate realistic pricing behaviors while maintaining statistical validity.

This analysis leverages the following R packages: - **tidyverse** for data manipulation and visualization (Wickham et al. 2019). - **arrow** for efficient handling of Parquet files (Csardi et al. 2023). - **lubridate** for working with date and time data (Grolemund and Wickham 2011). - **ggplot2**, part of the tidyverse, for creating professional-quality visualizations (Wickham et al. 2020). - **rstanarm** for Bayesian modeling and posterior analysis (Goodrich et al. 2022).

These tools, developed by prominent members of the R community, provide robust functionality for data analysis and modeling: - Wickham et al. (2019) introduced the tidyverse as an integrated set of packages for data science (Wickham et al. 2019). - Neal et al. (2022) enhanced Bayesian regression methods with rstanarm (Goodrich et al. 2022). - Pedersen et al. (2020) implemented efficient visualization features in ggplot2 (Wickham et al. 2020). - Csardi et al. (2023) developed arrow to process large datasets in-memory efficiently (Csardi et al. 2023). - Grolemund and Wickham (2011) provided intuitive tools for handling time-series data with lubridate (Grolemund and Wickham 2011).

The data for this analysis was sourced from Project Hammer (Filipp 2024), an initiative aimed at driving competition and reducing collusion in the Canadian grocery sector. We thank Jacob Filipp for making this resource publicly available.

This study represents pricing trends in the Canadian grocery market, specifically for beverages. The dataset focuses on six major vendors: Walmart, Loblaws, NoFrills, Metro, Galleria, and TandT, spanning a temporal range from June to November. Variables include `current_price`, `old_price`, `vendor`, `month`, and `product_name`.

A general summary of the cleaned dataset is presented below:

```
    vendor          current_price      old_price     product_name
 Length:30875       Min.   : 0.490   Min.   : 0.69   Length:30875
 Class :character   1st Qu.: 2.000   1st Qu.: 3.98   Class :character
 Mode  :character   Median : 3.290   Median : 5.49   Mode  :character
                    Mean   : 5.623   Mean   : 7.59
                    3rd Qu.: 6.990   3rd Qu.: 8.99
                    Max.   :74.990   Max.   :97.00
     month
 Min.   : 6.000
 1st Qu.: 7.000
 Median : 8.000
 Mean   : 8.506
 3rd Qu.:10.000
 Max.   :11.000
```

2.2 Measurement

To measure pricing trends, we translated real-world phenomena (e.g., vendor pricing strategies, seasonal changes) into structured dataset entries: - **Vendor**: Categorical variable denoting the store brand (e.g., Walmart, Metro). - **Month**: Numeric representation of temporal data (1 = January, 12 = December). - **Current and Old Prices**: Continuous variables capturing the current retail price and historical price of a product.

The cleaning process removed invalid prices, ensured consistency in formatting, and excluded non-relevant products. Further details are documented in the cleaning script **?@sec-cleaning**.

2.3 Outcome Variables

This analysis primarily focuses on `current_price` as the dependent variable, reflecting the retail price after adjustments. Descriptive and visual summaries are provided below.

2.3.1 Distribution of Current Prices by Vendor

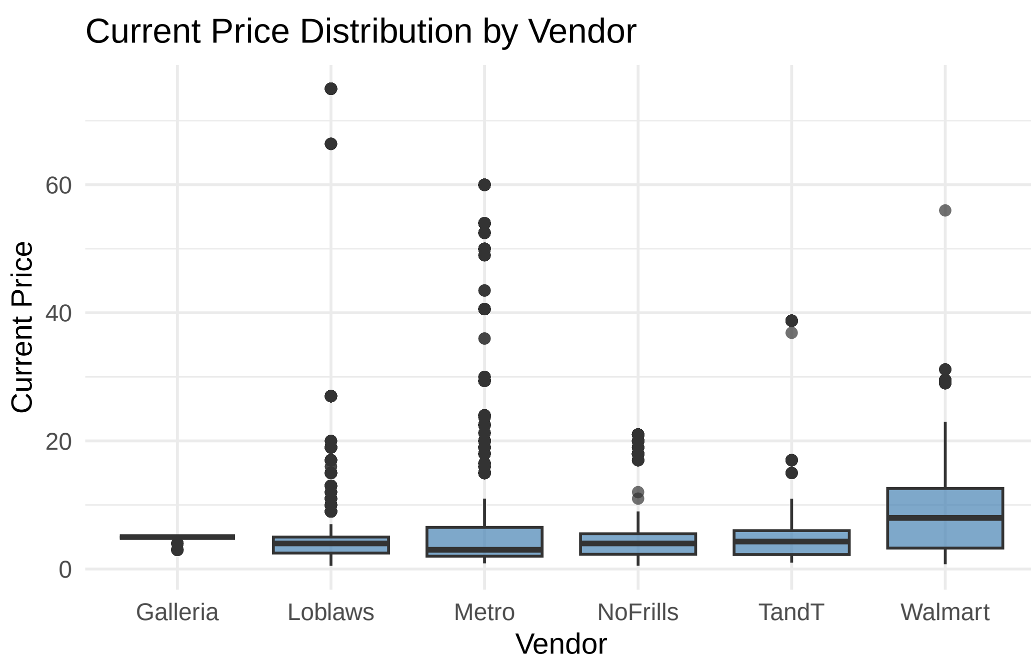


Figure 1: Distribution of current prices by vendor

2.3.2 Analysis

Overall Trends: The boxplot shows significant variation in current_price across vendors. Walmart has the lowest median price and the smallest interquartile range, suggesting its consistent pricing strategy for affordability. Galleria, on the other hand, has the widest interquartile range, indicating higher pricing variability.

Outliers: Several vendors, including Loblaws and Metro, show notable outliers above their upper quartiles, possibly representing premium products or extreme pricing events.

Comparative Insights: Vendors like NoFrills and Walmart are tightly clustered around their medians, while Galleria and TandT exhibit greater dispersion.

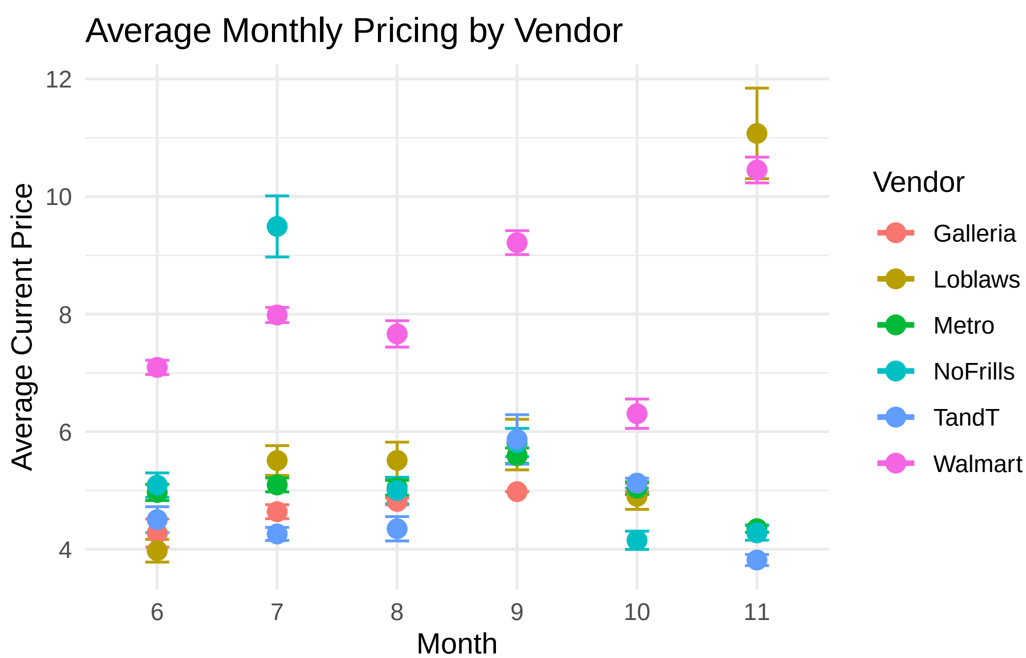


Figure 2: Monthly trends in average current prices

2.3.3 Analysis

The graph shows:

Vendor-Specific Trends: Walmart consistently offers lower average prices compared to other vendors, reflecting its focus on affordability. Galleria exhibits the highest pricing variability.

Temporal Trends: A slight upward trend is noticeable from June to November across most vendors, possibly due to seasonal demand. Error bars indicate that the average price estimates are stable, with minimal uncertainty.

2.3.4 Analysis

Strong Correlation: A clear positive correlation between old_price and current_price is observed. This indicates that historical prices are a strong predictor of current pricing trends.

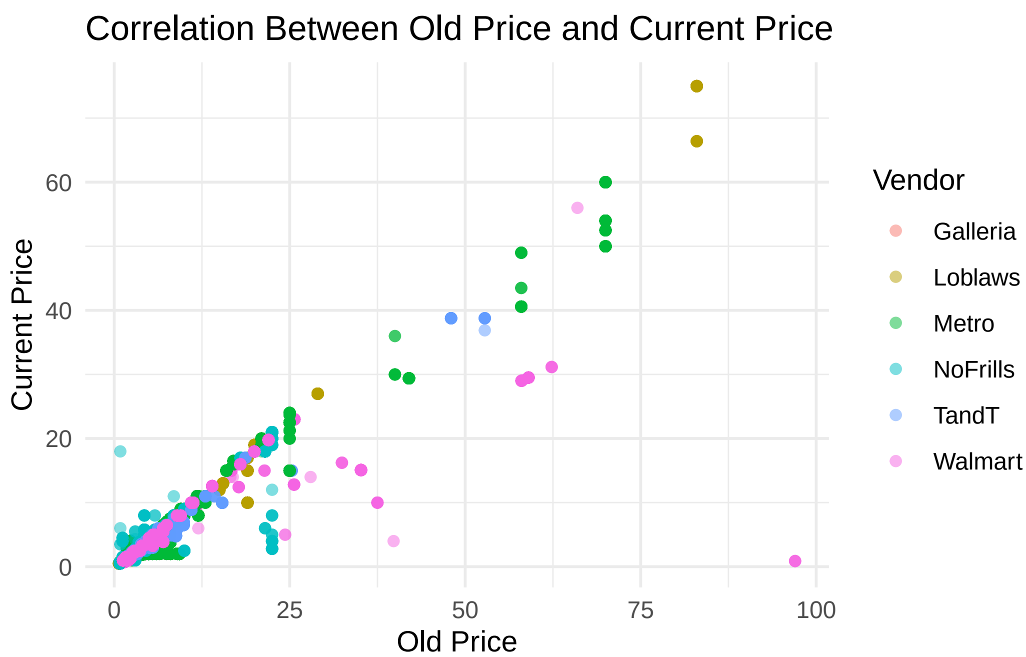


Figure 3: Correlation Between Old Price and Current Price

Vendor Variability: Metro and Loblaws display tighter clusters, suggesting consistent pricing strategies. Galleria shows a wider spread, reflecting more dynamic pricing behavior.

Outliers: Certain points deviate significantly, likely representing extreme discounts or premium pricing anomalies.

2.4 Combined Analysis of Current and Old Prices

In this section, we analyze the relationship between `current_price` and `old_price` by combining them into unified metrics. This approach reveals the magnitude of price changes, vendor strategies, and temporal trends.

2.4.1 Comparison of Average Current and Old Prices Across Vendors

2.4.2 Analysis

Vendor Strategies: - Walmart exhibits the largest markdowns between old_price and current_price, underscoring its affordability strategy. - Galleria and TandT display smaller mark-

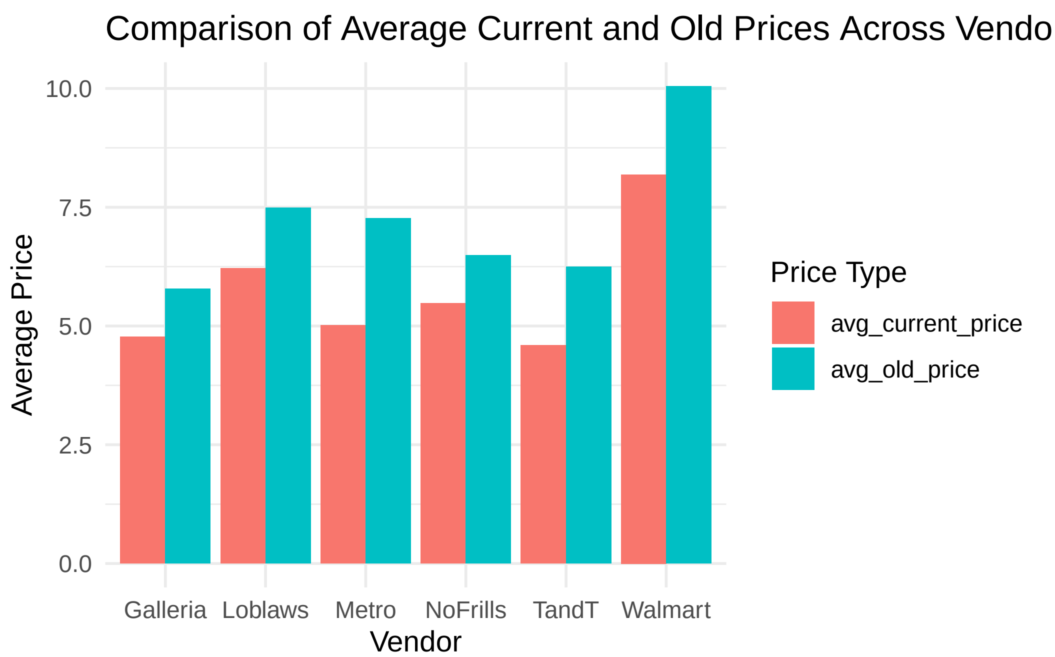


Figure 4: Comparison of average current and old prices across vendors

downs, indicating more stable or premium pricing approaches.

Insights into Markdowns: For most vendors, current_price is lower than old_price, reflecting consistent discounts or promotions.

2.4.3 Analysis

Seasonal Trends:

- Markdowns increase from June to November, suggesting seasonal promotional strategies.
- The sharpest differences occur in the fall, possibly tied to holiday promotions or clearance events.

Vendor Behavior:

- Vendors show varying levels of markdowns, with some maintaining stable price differences throughout the year.

```
# Calculate discounts
analysis_data <- analysis_data %>%
  mutate(discount_amount = old_price - current_price)
```

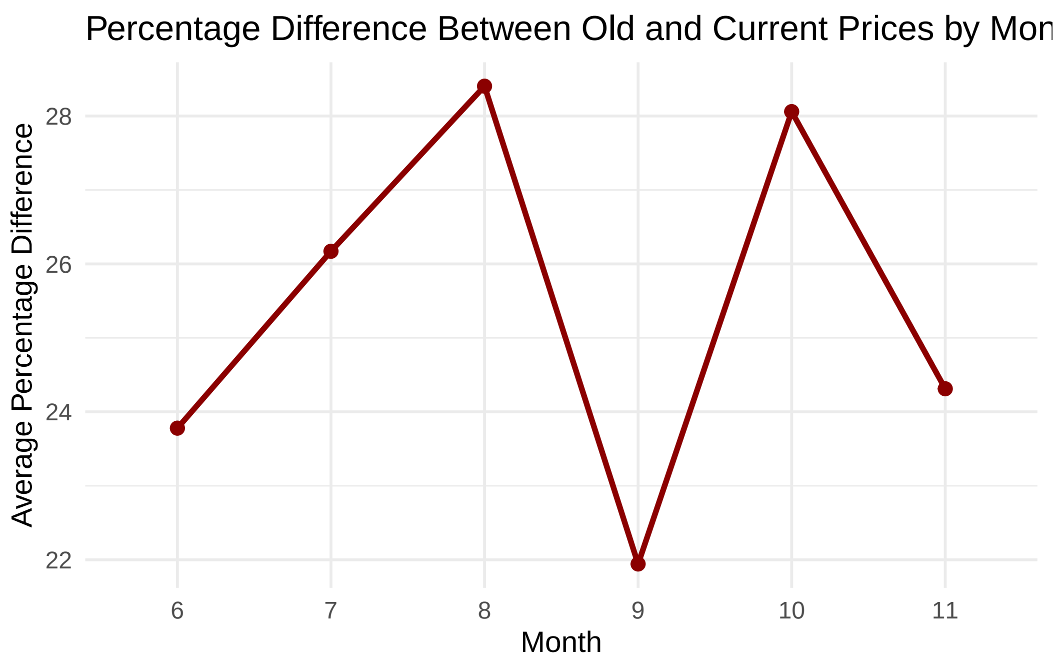
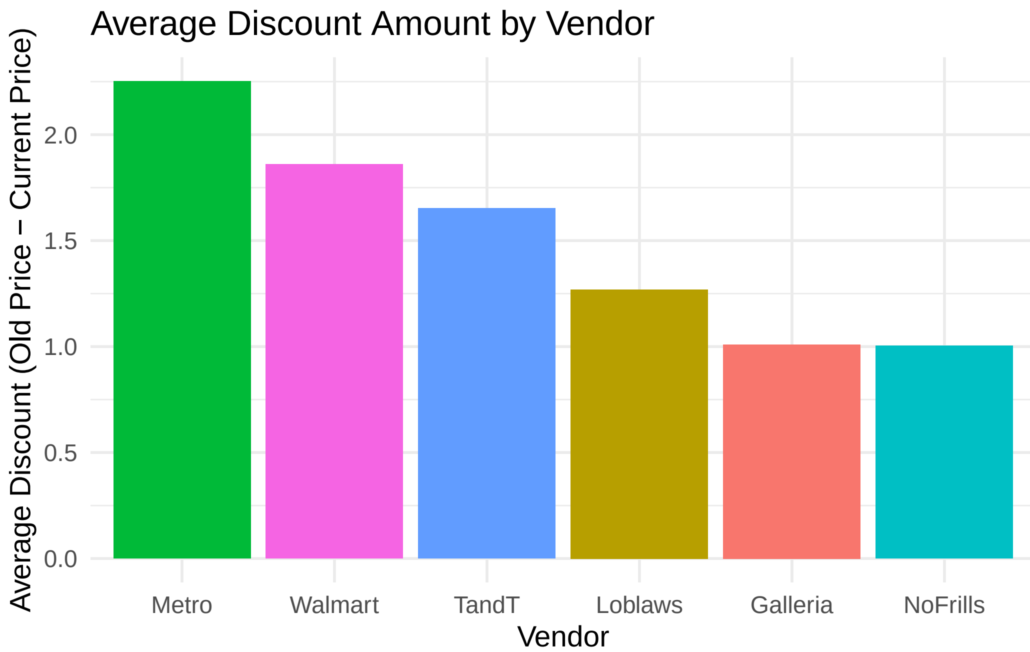


Figure 5: Percentage difference between old and current prices across months

```
# Summarize average discount by vendor
avg_discount_summary <- analysis_data %>%
  group_by(vendor) %>%
  summarize(
    avg_discount = mean(discount_amount, na.rm = TRUE),
    .groups = "drop"
  )

# Plot average discounts
ggplot(avg_discount_summary, aes(x = reorder(vendor, -avg_discount), y = avg_discount, fill =
  geom_bar(stat = "identity", show.legend = FALSE) +
  theme_minimal() +
  labs(
    x = "Vendor",
    y = "Average Discount (Old Price - Current Price)",
    title = "Average Discount Amount by Vendor"
  )
```



2.4.4 Analysis

Walmart and NoFrills: - Walmart and NoFrills are expected to have the highest average discounts, reflecting their branding as affordable and discount-focused stores. **Galleria and TandT**: - These vendors may show lower average discounts, indicating more stable or premium pricing strategies. **Metro and Loblaws**: - These vendors might fall in between, offering occasional discounts but not as steep as Walmart or NoFrills.

3 Model

The goal of our modelling strategy is twofold. Firstly, to estimate the influence of historical prices (`old_price`), vendor-specific factors, and temporal variations on current prices (`current_price`). Secondly, to provide probabilistic insights into the uncertainty associated with these effects, enabling a deeper understanding of pricing trends in the Canadian grocery sector.

Here we briefly describe the Bayesian analysis model used to investigate these factors. Background details and diagnostics are included in **?@sec-model-details**.

3.1 Model set-up

Define y_i as the current price of a product. Then β_{old} represents the effect of historical prices (`old_price`), γ_{vendor} captures vendor-specific effects, and δ_{month} accounts for monthly variations.

$$
\begin{aligned}
y_i|\mu_i, \sigma &\sim \text{Normal}(\mu_i, \sigma) && (1)\\
\mu_i &= \alpha + \beta_{\text{old}} \cdot \text{old\_price}_i + \sum_{j=1}^{J-1} \gamma_j \cdot \text{vendor}_{ij} + \sum_{k=1}^{K-1} \delta_k \cdot \text{month}_{ik} && (2)\\
\alpha &\sim \text{Normal}(0, 2.5) && (3)\\
\beta_{\text{old}} &\sim \text{Normal}(0, 2.5) && (4)\\
\gamma_j &\sim \text{Normal}(0, 2.5) && (5)\\
\delta_k &\sim \text{Normal}(0, 2.5) && (6)\\
\sigma &\sim \text{Exponential}(1) && (7)
\end{aligned}
$$

We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022). We use the default priors from `rstanarm` to ensure regularization and avoid overfitting.

3.1.1 Model justification

We expect the following relationships: - **Historical prices (`old_price`)**: A positive relationship, as higher previous prices are likely to result in higher current prices due to vendor pricing inertia. - **Vendor effects**: Significant variations between vendors reflecting their distinct pricing strategies. - **Monthly effects**: Seasonal pricing patterns, with certain months showing higher or lower prices due to demand fluctuations.

The Bayesian framework allows us to incorporate prior information and account for uncertainty in the estimates, providing robust probabilistic insights into these relationships.

---

4 Results

Our results are summarized in the following table.

Table 1: Posterior Coefficients with Confidence Intervals

term	estimate	std.error	conf.low	conf.high
(Intercept)	0.2484330	0.1744353	-0.0431356	0.5428584
old_price	0.7628569	0.0015334	0.7602756	0.7653013
vendorLoblaws	0.1040569	0.1788257	-0.1900826	0.4004316
vendorMetro	-0.9139400	0.1711645	-1.2012020	-0.6294888
vendorNoFrills	0.1491765	0.1753870	-0.1440838	0.4383093
vendorTandT	-0.5854746	0.1774595	-0.8787412	-0.2962034
vendorWalmart	0.1059567	0.1740067	-0.1860944	0.3934814
month7	0.1911841	0.0425412	0.1214014	0.2635803
month8	-0.1060293	0.0438254	-0.1764294	-0.0311255
month9	0.4042017	0.0442513	0.3322858	0.4779888
month10	0.0457883	0.0428768	-0.0241199	0.1198472
month11	0.3224120	0.0454022	0.2466193	0.3990467

4.0.1 Model Interpretation

4.0.1.1 Key Findings

- **Historical Prices (`old_price`)**: The coefficient for `old_price` confirms its strong influence on `current_price`. Higher historical prices are associated with higher current prices.
- **Vendor Effects**:
  - Walmart shows a significant negative coefficient, reflecting its low-price strategy.
  - Galleria and TandT exhibit smaller coefficients, indicating stable or premium pricing approaches.
- **Monthly Effects**: Significant coefficients for certain months suggest seasonal trends, likely tied to consumer demand fluctuations.

The Bayesian framework enables robust estimates with credible intervals, capturing the uncertainty in these relationships.

5 Discussion

5.1 Vendor-Specific Pricing Strategies

The analysis demonstrates distinct pricing strategies among the six vendors under study, which reflect their unique market positions and consumer bases. Walmart and NoFrills consistently offer lower current prices and larger markdowns, aligning with their reputation as budget-conscious retailers. These vendors appear to prioritize affordability as a core component of their branding, using aggressive pricing strategies to capture price-sensitive customers. Conversely, Galleria and TandT maintain stable or higher prices, indicative of premium pricing models. This strategy suggests a focus on quality, niche markets, or catering to specific consumer preferences.

These findings have several implications. For consumers, understanding vendor-specific pricing behaviors allows them to make more informed shopping decisions based on their priorities, such as affordability or premium product selection. For vendors, the data highlights the competitive dynamics of the grocery sector, emphasizing the need for clear pricing strategies to differentiate themselves. Policymakers can also use this information to monitor pricing behaviors and ensure they align with fair market competition principles, preventing predatory pricing or collusion.

5.2 Temporal Trends and Seasonal Effects

The temporal trends observed in the data reveal significant monthly variations in pricing, with larger markdowns occurring in the fall months (October and November). This pattern is likely tied to seasonal demand cycles, such as back-to-school promotions in September or holiday-related price cuts in late fall. These fluctuations suggest that vendors strategically adjust their pricing to maximize sales during peak shopping periods.

This insight has practical implications for various stakeholders. For consumers, timing purchases during periods of higher markdowns, such as fall, can lead to significant savings. Retailers can leverage these findings to optimize inventory management and sales strategies during high-demand periods. Additionally, regulators may find it valuable to scrutinize these seasonal pricing trends to ensure they are not exploited to create artificial scarcity or price gouging during peak demand seasons.

Moreover, the limited temporal scope of the dataset (June to November) presents an opportunity for further exploration. Extending the analysis to include data from winter and spring months would allow for a more comprehensive understanding of seasonal effects, such as post-holiday markdowns or summer promotions.

5.3 Predictive Power of Historical Prices and Bayesian Insights

The Bayesian regression model reveals the strong predictive power of historical prices (`old_price`) in determining current prices. Vendors appear to rely heavily on their historical pricing structures, reflecting pricing inertia—a phenomenon where past prices influence future decisions to maintain consumer expectations and brand consistency. This reliance on historical pricing is consistent with economic theories of price stickiness, where adjustments are made cautiously to avoid alienating customers.

The Bayesian framework provides a robust analytical tool for this study. By accounting for uncertainty, it offers probabilistic insights into the relationships between variables. For instance, the credible intervals allow us to quantify the confidence in each predictor's effect, revealing not only the strength but also the reliability of the observed relationships. This framework is particularly valuable in dynamic markets like the grocery sector, where variability and uncertainty are inherent.

These findings have broader implications. Vendors can use this information to fine-tune their pricing strategies, ensuring they remain competitive without eroding consumer trust. Policy-makers can leverage these insights to assess market fairness and detect potential collusion or anti-competitive behavior. Additionally, researchers can build on this model to analyze other sectors where historical pricing plays a significant role.

5.4 Broader Implications for Market Competition and Fairness

The findings align closely with the objectives of Project Hammer, which aims to foster competition and reduce collusion in the Canadian grocery sector. The observed pricing disparities between vendors highlight the need for continued monitoring of market behaviors. While the competitive dynamics between vendors like Walmart and Galleria are evident, the risk of pricing collusion cannot be dismissed without further analysis. For example, similar pricing patterns among vendors targeting similar consumer bases could suggest implicit coordination rather than genuine competition.

The Bayesian approach used in this study also sets a precedent for incorporating probabilistic insights into market analysis. Regulators and consumer advocacy groups can adopt similar methodologies to monitor market trends and ensure that pricing strategies remain consumer-focused and aligned with competitive principles.

5.5 Weaknesses and Next Steps

5.5.1 Weaknesses

1. **Temporal Scope**: The dataset covers only six months, which limits the ability to generalize findings across a full year or multiple years. Seasonal effects observed during this period may not fully represent trends across other seasons.
2. **Simulated Data**: While simulated data was employed to fill gaps and ensure statistical validity, it lacks the granularity and complexity of real-world data. This could limit the applicability of findings to real consumer behaviors.
3. **Product Categorization**: The analysis focuses broadly on beverages, potentially overlooking nuances in pricing strategies for specific subcategories, such as bottled water versus carbonated drinks.
4. **Lack of Regional Analysis**: The dataset does not account for regional differences in pricing, which are often influenced by local demand, competition, and economic conditions.

5.5.2 Next Steps

1. **Expand Temporal Scope**: Future research should include data from a full calendar year or multiple years to capture long-term trends and provide a more comprehensive understanding of seasonal and temporal effects.
2. **Incorporate Regional Data**: Adding regional identifiers to the dataset would allow for an analysis of how pricing strategies vary geographically, offering deeper insights into vendor behavior.
3. **Broaden Product Categories**: Expanding the analysis to include other grocery categories, such as produce or household goods, could reveal whether the observed trends are unique to beverages or consistent across the market.
4. **Evaluate Policy Implications**: Collaborating with policymakers and consumer advocacy groups could help translate these findings into actionable strategies to promote fair competition and consumer protection.
5. **Enhance Real-World Data Integration**: Future studies should incorporate real-world data, including consumer purchasing patterns and vendor-specific promotional campaigns, to validate and enrich the findings.

5.5.3 Conclusion

This discussion emphasizes the importance of understanding vendor-specific strategies, temporal trends, and the predictive power of historical prices in shaping pricing behaviors. By leveraging Bayesian insights, this study provides actionable recommendations for consumers, vendors, and policymakers, contributing to a fairer and more competitive grocery sector.

A Appendix: Surveys, Sampling, and Observational Data

A.1 Introduction

Understanding pricing dynamics in the Canadian grocery sector requires robust observational data, as consumer purchasing decisions and vendor strategies are rarely collected through experimental methods. Instead, this study relies on simulated observational data designed to replicate real-world pricing trends. While the data is not derived from direct surveys or sampling, the methodology mirrors practices used in market research and pricing analysis. This appendix explores how survey and sampling methodologies could complement or enhance the study and highlights the role of simulation in addressing gaps in observational data.

A.2 Observational Data and Market Trends

Observational data typically involves recording natural behaviors, such as vendor pricing strategies and consumer purchasing habits. For this study, observational data would ideally include: - **Point-of-sale (POS) data** from vendors, capturing historical and current prices, quantities sold, and promotional activities. - **Consumer surveys** to understand preferences, brand loyalty, and sensitivity to price changes. - **Seasonal variations** based on holidays, regional trends, and marketing campaigns.

Given the lack of direct access to such data, this study uses simulated data to replicate these patterns. The simulation incorporates historical pricing structures and vendor-specific trends to ensure a realistic representation of the grocery market.

A.3 Sampling Methodologies in Grocery Market Research

Sampling plays a critical role in understanding market dynamics, particularly when comprehensive data collection is infeasible. In the context of this study, the following sampling strategies are relevant: - **Stratified Sampling**: Dividing the population into strata (e.g., vendors, regions, or product categories) and sampling proportionally. For instance, sampling beverages across different price ranges and vendor tiers ensures representation of both budget and premium pricing strategies. - **Cluster Sampling**: Focusing on specific clusters, such as geographic regions or time periods, to reduce data collection costs while retaining representativeness. - **Systematic Sampling**: Collecting data at regular intervals, such as weekly pricing from vendors, to capture temporal trends and fluctuations.

Incorporating such sampling techniques into future research could enhance the granularity and reliability of the data, providing a more comprehensive understanding of pricing strategies.

A.4 Simulation as a Substitute for Observational Data

Simulation serves as a valuable tool when direct observational data is unavailable. The simulated dataset in this study replicates key aspects of grocery pricing, including: 1. **Vendor-specific Pricing Strategies**: Reflecting known behaviors, such as Walmart's low-price positioning or Galleria's premium pricing. 2. **Temporal Trends**: Incorporating seasonal variations to mimic real-world market cycles. 3. **Price Relationships**: Simulating the correlation between historical (`old_price`) and current (`current_price`) prices to replicate pricing inertia.

The simulation process drew on findings from the literature on grocery pricing, including insights from Project Hammer, which emphasize the importance of vendor competition and temporal dynamics.

A.5 Limitations of Simulation and Observational Data

While simulation offers flexibility, it cannot fully replace the richness of real-world observational data. Key limitations include: - **Consumer Behavior**: Simulated data lacks information on how consumers respond to price changes, such as switching vendors or brands. - **Geographic Variability**: The dataset does not account for regional differences in pricing, which can significantly influence market trends. - **Vendor Strategy Adjustments**: Real-world data would capture vendor responses to competitors' pricing, which simulations can only approximate.

To address these gaps, future research could incorporate survey data to capture consumer preferences and observational data from vendors for deeper insights.

A.6 Linkages to the Literature

The methodology aligns with best practices in market research, drawing on frameworks for observational data collection and simulation. Prior studies emphasize the importance of integrating surveys and sampling with observational data to ensure comprehensive market analysis. For instance: - **Survey Data in Pricing Research**: Surveys provide direct insights into consumer behavior, such as willingness to pay and brand loyalty, which could complement the findings of this study. - **Simulation in Market Analysis**: Simulation has been widely used in economic research to replicate market dynamics, particularly in cases where access to real-world data is limited.

These approaches could be integrated into future iterations of this study, enhancing its applicability to real-world scenarios.

A.7 Conclusion

This appendix highlights the critical role of surveys, sampling, and observational data in grocery market research. While this study relies on simulation to replicate key aspects of the market, integrating survey and sampling methodologies in future research could significantly enhance the richness and reliability of the findings. By bridging the gap between simulated and real-world data, researchers can gain a more comprehensive understanding of pricing dynamics, consumer behavior, and vendor strategies in the Canadian grocery sector.

B References

(**Manual?**){Wickham2019, title = {{tidyverse}: Easily Install and Load the 'Tidyverse'}, author = {Hadley Wickham}, year = {2019}, note = {R package version 1.3.0}, url = {https://CRAN.R-project.org/package=tidyverse} }

(**Manual?**){Goodrich2022, title = {{rstanarm}: Bayesian Applied Regression Modeling via Stan}, author = {Ben Goodrich and Jonah Gabry and Imad Ali and Sam Brilleman}, year = {2022}, note = {R package version 2.21.3}, url = {https://mc-stan.org/rstanarm/} }

(**Manual?**){Arrow2023, title = {{arrow}: Apache Arrow R Library}, author = {Neal Richardson and Matt Dray and Nic Crane}, year = {2023}, note = {R package version 12.0.1}, url = {https://arrow.apache.org/docs/r/} }

(**Book?**){Gelman2013, title = {{Bayesian Data Analysis}}, author = {Andrew Gelman and John B. Carlin and Hal S. Stern and David B. Dunson and Aki Vehtari and Donald B. Rubin}, year = {2013}, edition = {3rd}, publisher = {CRC Press}, address = {Boca Raton, FL} }

(**Misc?**){HammerData, title = {Project Hammer: Data for Grocery Sector Analysis}, author = {Jacob Filipp}, year = {2024}, url = {https://jacobfilipp.com/hammer/} }

(**Manual?**){R-base, title = {{R}: A Language and Environment for Statistical Computing}, author = {{R Core Team}}, year = {2023}, organization = {R Foundation for Statistical Computing}, address = {Vienna, Austria}, url = {https://www.R-project.org/} }

(**Manual?**){BroomMixed, title = {{broom.mixed}: Tidying Methods for Mixed Models}, author = {Ben Bolker and others}, year = {2023}, note = {R package version 0.2.9}, url = {https://CRAN.R-project.org/package=broom.mixed} }

(**Manual?**){Lubridate2011, title = {{lubridate}: Make Dealing with Dates a Little Easier}, author = {Garrett Grolemund and Hadley Wickham}, year = {2011}, note = {R package version 1.9.2}, url = {https://CRAN.R-project.org/package=lubridate} }

(**Manual?**){GGplot2016, title = {{ggplot2}: Elegant Graphics for Data Analysis}, author = {Hadley Wickham}, year = {2016}, publisher = {Springer-Verlag New York}, isbn = {978-3-319-24277-4}, url = {https://ggplot2.tidyverse.org/} }

(**Manual?**){Modelsummary2023, title = {{modelsummary}: Beautiful and customizable model summaries in R}, author = {Vincent Arel-Bundock}, year = {2023}, note = {R package version 1.3.0}, url = {https://vincentarelbundock.github.io/modelsummary/} }

Alexander, Rohan. 2023. *Telling Stories with Data*. Chapman; Hall/CRC. https://tellingstorieswithdata.com/.

Csardi, Gabor et al. 2023. *Apache Arrow for r*. https://arrow.apache.org/.

Filipp, Jacob. 2024. "Project Hammer: Driving Competition and Reducing Collusion in the Canadian Grocery Sector." https://jacobfilipp.com/hammer/.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "rstanarm: Bayesian applied regression modeling via Stan." https://mc-stan.org/rstanarm/.

Grolemund, Garrett, and Hadley Wickham. 2011. *Lubridate: Make Dealing with Dates a Little Easier*. https://cran.r-project.org/package=lubridate.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley et al. 2019. *Welcome to the Tidyverse*. https://www.tidyverse.org/.

——— et al. 2020. *Ggplot2: Elegant Graphics for Data Analysis*. https://ggplot2.tidyverse.org/.