December 1, 2023

To: Rachel Rosen, Chief Analytics Officer, Commonwealth Foods Inc.

From: Data Science Team - Andrew Barros, Mariska Batavia, Shrikant Mishra, Matt Scheffel

Subject: CFI Coupon Campaigns and Future Directions

Dear Rachel,

We have been working hard in preparation for the upcoming executive retreat and are looking forward to presenting it to the leadership team. As you know, we have focused our efforts on manufacturer coupons, and we think our findings will give a head start to the marketing team and will also persuade CFI's executives to invest in growing the data science team. For clarity, the memo below is organized into four parts: 1) a summary of the business problem, 2) the details of our analysis, 3) our conclusions and recommendations, and 4) the future of the data science team at CFI.

What is the business problem or opportunity and why does it matter to the business?

Having expanded to 580 stores, Commonwealth Foods, Inc (CFI) has enjoyed a recent wave of growth due to the popularity of its organic products amongst consumers. It has responded to this change by adopting several new practices, such as a frequent shopper program and the creation of private label products, as well as keeping its "local" feel with a manager-centric operating model. Recently, the company has made some changes in executive leadership - including bringing in a data-driven CEO and establishing a Chief Analytics Officer. Now, CFI wants to leverage a new data science team and emerging data science practices to enhance their business strategy and operations.

The primary objective of the newly formed data science team is to leverage the existing data to gain a deeper insight into the effectiveness of two major marketing expenditures at CFI: promotional activities (such as coupons) and the frequent shopper program. CFI utilizes a range of direct marketing strategies, including email campaigns, personalized deals, and focused promotions, to engage directly with customers. While these initiatives have been promising, it is critical to evaluate their real influence on factors such as customer loyalty, repeat buying behavior, and overall sales. Analyzing this data and gaining clarity on these aspects will allow CFI executives to guide future strategic choices and recommendations for marketing efforts, particularly in the areas of coupon distribution and identifying the most suitable customer segments for these campaigns.

What did you analyze and why? How did you go about it?

Original Analysis:

Our team embarked on a comprehensive analysis to evaluate the impact of coupon redemption on household spending, with the aim of gauging the effectiveness of the company's promotional coupon strategies. We were presented with detailed data on a cohort of 2,500 households who are frequent shoppers at CFI. The data had detailed records of all products purchased, information on the marketing campaigns received by each household, and additional details about the stores and products. Additionally, there was demographic data available for approximately 30% of the cohort.

The first and most intensive task for our team required transforming the raw data into a format which was suitable for our analysis. The process began with the identification of coupon redemption events, which occur when a household redeems a coupon. We compiled a comprehensive list of products that could be purchased with each coupon. Then, for each redemption event, we collated and summarized the household's spending activities leading up to, on the day of, and following the event, focusing on total expenditure, spending on potentially covered products, and the number of transactions. Finally, we normalized the spending data on a per-transaction basis to facilitate uniform analysis.

Originally, our team sought to use the transaction records to identify both product-specific and total spending before and after the receipt of a coupon. For the subset with available demographic data, we also sought to determine if the effect of the coupon was impacted by age, household size, or reported income. However, during the preparation and exploratory data analysis phase, our team encountered notable limitations, such as an incomplete linkage between products and coupons due to the complexity of multiple products being associated with single coupons, and vice versa. Moreover, the provided date information was inadequate for analyzing seasonal spending trends, presenting a significant hurdle in understanding time-bound consumer behaviors.

Adjustments and Further Discoveries:

Adapting to these challenges, our team proceeded with a preliminary analysis using a mixed-effects linear model. This initial foray revealed that while coupon redemption did not significantly affect product-specific spending, it was associated with an average increase of $1.50 in total future spend per transaction. These insights were derived from a detailed examination of three predominant campaigns which together accounted for a substantial 75% of all redemption events. Notably, we determined that these campaigns were likely driven by a repeat marketing effort by a singular national manufacturer and exhibited considerable overlap in the “Frozen Meat/Meat Dinners” category.

A deeper dive into the demographic data revealed a homogeneity across the campaigns, indicating that customer demographics did not markedly affect spending patterns post-coupon redemption. For transactions where demographic data were available, the demographics of

customers did not differ noticeably between the campaigns: the highest proportion of customers were 45-54 years old (approximately 43%), had household sizes of 2 or 1 (approximately 68%), and made $50-74,000 per year (approximately 35%). Household size in coupon users was comparable to the general customer base overall, but coupon users were overrepresented in the 45-54 age group and $50-74,000 income bracket. While these demographic groups appear more likely to redeem coupons, it also appears that none of the demographic factors considered - age, household size, and income - significantly impact future spending after redeeming a coupon. Additionally, out of the cohort of 2,500 households that we had data for, demographic information existed for only 801 households. Furthermore, out of those households, only 301 had redeemed coupons at CFI. While this did not stop us from answering our original research questions, the reduced number of households in the analysis may affect the strength of any recommendations we give on coupon use in regard to demographic data.

Final Steps:

To refine the understanding of customer behavior further and based on some empirical observations of the data, our team employed unsupervised clustering techniques on the demographic data, revealing that families with children were the most likely to redeem coupons. Across all clustering scenarios, both income and age were ranked low in feature importance, and household size/composition and marital status had higher importance. The highest silhouette score was for 3 clusters (0.180). The first cluster consisted mainly of couples with no children; the second cluster consisted of married couples with one or more children; and the final cluster included primarily single adults living alone. Married couples with children accounted for 42% of transactions with coupons, and only 31% of transactions without coupons. Married couples without children accounted for 33% of transactions with coupons and 31% of transactions without coupons. Single adults living alone accounted for 26% of transactions with coupons and 33% of transactions without coupons.

From this segmentation analysis, it is clear that the demographic which is most likely to take advantage of coupons is families with children. If CFI's management team deems coupons to be a valuable investment, that demographic group could be a prime target for more focused marketing efforts. Intriguingly, the analysis also indicated that the average coupon discount per transaction varied across demographic clusters, with married couples without children receiving the highest discounts ($1.45) and families with children having the lowest mean discount per transaction ($1.12). Our team interpreted this to mean that families with children may be willing to use coupons even if the discount is small, whereas married couples without children and single adults may only be willing to use higher value coupons.

Our analysis culminated in the further development and use of models which estimated the differential impact on total and product-specific spending due to coupon redemption. These models consistently showed an increase in spending across various customer segments and were particularly indicative of established customers' behavior.

What did you learn and what do you recommend to the business?

One of Z's next goals was to determine the efficacy of promotions for CFI. Through our analysis, we specifically focused on the effectiveness of coupon promotions. Here is a summary of our learnings:

- Coupon redemptions are associated with an overall increase in total future spend per transaction, but there is no visible impact on product-specific spend
- Campaigns 8, 13, and 18 combine for most coupon redemptions, and have substantial overlap in the products covered
- No specific demographic factor significantly impacts future spending
- Married couples with children accounted for the most coupon transactions out of all demographic clusters
- Single adults and married couples without children have higher average coupon discounts per transaction

Keeping these points in mind, we would like to recommend to Z that **coupon promotions are effective.** While there is no visible impact on product-specific spend, a $1.50 in total future spend per transaction does show that customers tend to spend more in post-coupon-redemption transactions. Coupon discounts are applied, and in many cases, discovered, at checkout lanes. The revelation that products are going to have marked-down prices could serve as positive reinforcement for customers to return and spend more on future visits. Additionally, while this value may appear small, extrapolating over thousands of customers, several months, and over 580 CFI stores shows that coupon discounts could lead to an increase in future sales by thousands of dollars overall. In this way, the coupon promotions are having the intended impact on our customers' spending habits and our future revenue. Finally, considering that we are primarily dealing with manufacturer's coupons, which are reimbursed by manufacturers upon redemption, keeping coupon promotions has little risk to our profits.

Next, we would also recommend that our coupon campaigns be adjusted so that they appeal to the demographic group that resulted in the highest amount of coupon transactions: families with children. Married couples with children accounted for 42% of transactions with coupons, and only 31% of transactions without coupons. This differential suggests that this demographic primarily prefers shopping for products marked with discounts. Additionally, the segmentation analysis described earlier showed that families with children had the lowest mean discount per transaction out of all three clusters. Combined with the earlier fact, this suggests that this demographic group is more interested in coupon promotions in general, regardless of the size of the discount. Perhaps receiving any type of price deduction on grocery products is a priority in families where expenses are higher due to the larger number of dependents. To maximize our sales, it would make sense to market flyers or send mailers with coupon promotions to such

households, or to partner with manufacturers that produce items that appeal to them. This could include common family products such as cereal.

What else should CFI do and how can the Data Science team continue to help with the business problem or opportunity?

Regarding our marketing efforts, and this business opportunity specifically, the primary action that CFI should take is to invest in expanding the amount of data that is available. From our experiences, we noted the following issues with the data:

- Transaction day fields cannot be anchored to a calendar day.
- Out of a total of 2500 households, demographics information is only available for 801 households. Out of these, only 301 had redeemed coupons.

First, we noticed that date information in the data referred to number of days from an arbitrary start date. We were unable to anchor the days to specific calendar dates, which would have allowed us to account for seasonality in our analysis. Our first recommendation would be to connect with Roy and update the data documentation to provide more context into the conversions for both date and time fields. With this data, the Data Science team can investigate customer spending behavior, through the perspective of coupon redemptions, over the course of several months and years. We expect that coupons are more likely to be redeemed during certain periods of the year, such as the holiday season, where food expenses are generally higher, or when certain fruits are "in season." Access to such data would enable us to confirm our hypothesis. Additionally, with interpretable time fields as well, we could have conducted similar analyses for hours over a business day i.e., whether coupons are more likely to be redeemed during certain hours of the day.

Second, since only 32% of households in the dataset had associated demographic data available, we would have preferred a higher proportion to increase the strength of our recommendation. Understandably, due to the personal nature of the data, we expect that demographic data would not be available for all of the households in the dataset. However, CFI could alleviate this issue by exploring ways to make the process more convenient and accessible for shoppers. If demographic information was obtained via survey, successful completion of one could be rewarded with a future coupon to further incentivize customers.

As the Data Science team at CFI, we can use the earlier proposed changes to create even more robust recommendations on coupon promotions. Additionally, we can build off our existing analysis to promote coupon promotions to other manufacturers that could partner with CFI. Manufacturers look to issue coupons as a form of product promotion. The discounted price certainly attracts the attention of shoppers who may not be considering the product otherwise, due to unaffordable prices, brand unfamiliarity, or disinterest. While we cannot prove that product-specific spend increases after coupon redemption, the Data Science team can investigate other avenues within the data for showcasing the benefits of having coupons to the

manufacturers themselves. This could also extend beyond coupon promotions to advertising potential product placement within our stores.

We hope this analysis makes sense and sparks future conversation about how the Data Science Team can further the business goals of CFI. We look forward to discussing this with you further.

Best,

Andrew Barros, Mariska Batavia, Shrikant Mishra, Matt Scheffel