Effects of Expressing Gratitude for Shopping Local via Sign on Tipping Behavior

Masters in Information and Data Science Class 241

Abby Purnell, Vicky Liu, Lia Cappellari, Lily Magliente

Abstract

In recent times, there has been a significant push to increase tipping culture and support local businesses. While many studies could be conducted to analyze the impact of these trends, this study instead focuses on actionable and attainable steps a local business can take to increase tips for their staff. This study explores the influence of expressing appreciation for supporting local businesses on consumer tipping behavior. In collaboration with MCM Coffee, a coffee shop in New York City, we conducted an experiment where a sign that reads "Thank you for supporting local" was placed next to the register on treatment days and removed on control days, and we compared the tipping behavior between the treatment and control groups. All customers who visited the coffee shop within the two week experiment period who paid with a credit or debit card were assigned to either the treatment (N = 848) or control group (N = 683) depending on the date of their purchase. The results suggested that on average, the treatment group tipped 1.08% higher than the control group, although this increase was not significant at the 5% level. Future experiments should consider increasing the sample size and controlling for additional covariates, such as the day of the week or the barista, to further investigate this causal relationship.

Introduction

In the past few years, there has been a growing movement towards supporting local business in the community rather than continuing to contribute to the growth and wealth of large corporations. At the same time, more and more fast casual restaurants and coffee shops have introduced new point-of-sale systems which automatically prompts customers to leave a tip. This trend presents new opportunities to reinforce the importance of supporting local businesses and potentially increase tips for staff. With this in mind, we aim to test the following: ***Does adding a sign that says "Thank you for supporting local!" at the register of a coffee shop increase tips?***

Our study is motivated by a survey conducted by Intuit in 2023, which revealed that 70% of consumers are actively supporting local businesses, and 82% are willing to spend more at these establishments post-pandemic to help sustain them[1]. Our treatment aims to remind consumers of the positive impact their support has on the local community, encouraging them to continue this behavior and potentially increase their tipping as a tangible expression of their appreciation. Additionally, our study builds on the idea of adding personalization in customer interactions, as demonstrated in previous research done by Michael Lynn from Cornell University (Lynn, 1996)[2]. He found that adding personalization from a server's side to provide service to customers, like introducing oneself by name, adding a "Thank you" and a smiley face on the check, yielded higher tips. Our study furthers this research to see if gently reminding customers they are supporting local business will change their tipping habits. This approach offers a simple, yet cost-effective strategy to boost earnings for local employees working in small businesses.

Hypotheses

The specific hypotheses for this experiment are:

Null Hypothesis (H_0): Displaying an appreciation sign does not increase the tipping percentage of customers at MCM Coffee Shop.

Alternative Hypothesis (H_1): Displaying an appreciation sign increases the tipping percentage of customers at MCM Coffee Shop.

We anticipate that the appreciation sign will lead to a statistically significant increase in tipping percentages based on the principle of reciprocity and the positive results observed in similar studies.

Treatment Design

The treatment involves placing a sign that reads "Thanks for supporting local" next to the payment area at MCM Coffee Shop. This sign serves as a non-monetary token of appreciation intended to influence customer tipping behavior. The control condition involves no sign being displayed.

[1] https://businessjournaldaily.com/70-of-americans-shop-small/

[2] Lynn, Michael. (1996). Seven Ways to Increase Servers' Tips. Cornell Hotel and Restaurant Administration Quarterly

Figure 1. Appreciation Sign

Participants in this study include all customers who made card transactions at MCM Coffee Shop during the two-week experiment period from June 24, 2024 to July 7, 2024. Cash transactions and tips made in the tip jar are excluded to ensure consistency in the data. The data set comprises transaction-level details from the shop's digital POS system, including total purchase amounts and tips for card transactions. Compliance with the treatment was monitored by ensuring the sign was displayed as planned on treatment days and removed on control days. The experimental design for a randomized two-group, posttest-only experiment can be denoted in the following ROXO diagram:

R X O

R O

The *Rs* in the diagram above denote the randomization technique - MCM Coffee's transactions were randomized into two groups, the control or treatment group, based on the day the transaction occurs.. The treatment is denoted by the *X*, and the potential outcomes of both groups are denoted by the *Os*. We compare the average tip percent between the control and treatment to understand the causality of our claim.

Potential Outcomes

The potential outcomes of interest are the average tipping percentages on days when the appreciation sign is displayed (treatment group) compared to days when the sign is not displayed (control group). Specifically, we aim to compare:

1. Treatment Group Outcome (Y_1): The average tipping percentage on days when the appreciation sign is displayed. This is calculated at the transactional level as the tip amount divided by the before-tax sales amount for each transaction.
2. Control Group Outcome (Y_0): The average tipping percentage on days when the appreciation sign is not displayed. This is calculated at the transactional level in the same manner as the treatment group.

The primary metric for comparison is the average tip percentage, which is defined at the transactional level. By comparing these average tip percentages between the treatment group (Y_1) and the control group (Y_0), we aim to determine the impact of the appreciation sign on customer tipping behavior. This difference (Y_1 - Y_0) represents the causal effect of the appreciation sign on tipping behavior.

Randomization

To ensure unbiased results, we aimed to implement a randomized controlled trial (RCT) with a balanced randomization process. The experiment was conducted over two weeks, starting June 24, 2024, where we provided the coffee shop with a schedule where treatment and control days alternated. This approach ensures that each day of the week is equally represented in both the treatment and control conditions, thereby accounting for any day-specific variations in tipping behavior.

Although alternating days may not be strictly random, we aimed for a practical and systematic approach to ensuring that external factors such as day-of-week effects were evenly distributed across treatment and control conditions. This method helps mitigate potential biases that could arise from non-random assignment.

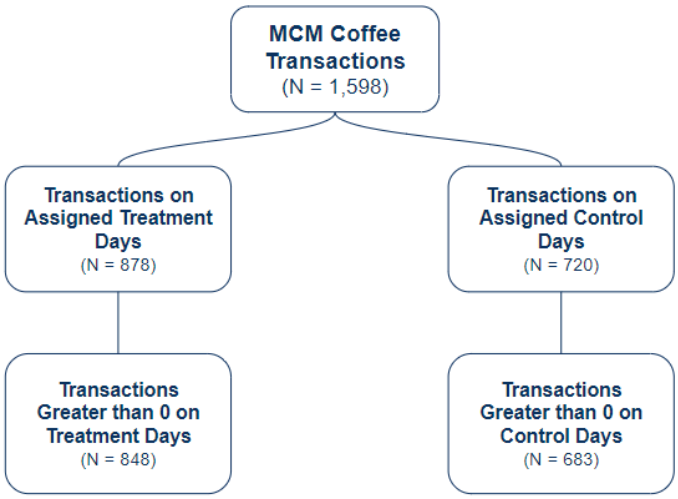

Figure 2. Flow chart of Treatment versus Control Assignments

After the study was conducted and we had collected the data, we determined the coffee shop was closed on the 4th of July holiday, leaving us with one less control day than treatment, potentially biasing our results. We also determined there were 66 transactions included in the report that had $0 of total sales, tax, or tip. We removed those data points from our dataset as we believe those

may have been employee transactions or transactions that were voided, and therefore should not be considered in our analysis.

Initial Power Analysis

Based on initial conversations with MCM Coffee Shop, we estimated an average of 150 customers per weekday and 100 customers per weekend day, totaling approximately 1,700 customers over the two-week period. Our pre-experiment power analysis indicated that with a sample size of 1,000 customers, we could achieve 80% power to detect a small effect size. This calculation assumed a baseline tipping percentage of 6% and a potential increase to 7% with the treatment, with standard deviations of 4% and 5%, respectively. These assumptions were based on preliminary data and informed by similar studies in the literature.

Exploratory Data Analysis

We ran our experiment for two weeks, where Treatment and Control Days alternated. We began on June 24th, 2024 with a Control Day and ended on July 7th, 2024 as a Treatment day. The only derivation from the initially planned schedule is that MCM Coffee was closed on the holiday, July 4th.

Below we see some summary statistics between Treatment and Control groups. There are 1,531 total units assessed in our experiment. The average tip amount on control days is $0.78 on Control days and $0.90 on Treatment days. 45% of the total customers are in the Control Group and 55% in the Treatment Group. The average tip amount in Control is 9.1% and Treatment is 10.1%. We have 361 people who tipped on Control days (out of 683 units) and 482 people who tipped (out of 848 total units). (Summary table below)

Treatment	Average Tip ($)	Tip Standard Deviation ($)	Average Tip Percent (%)	Tip Percent Standard Deviation (%)	Number of Tippers	Number of Customers	Percent of Transactions
Control	0.780	1.026	0.091	0.102	361	683	0.446
Treatment	0.902	1.097	0.101	0.104	482	848	0.554

Figure 3. Summary Statistics for Treatment vs Control

Figure 4 below initially suggests that there are more tips in total dollars tipped on Treatment Days as compared to Control Days, but also notice that there are more Sales in Dollars on Treatment days than compared to Control days.

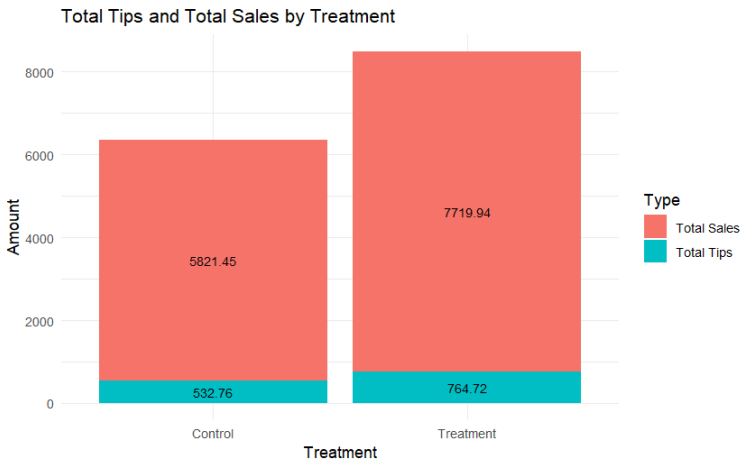

Figure 4. Total Tips and Sales in Dollars for Treatment vs Control

Figure 5 visualizes the distribution of Percentage Tipped Before Tax by Treatment. For both Treatment and Control days, both the minimum and 1st quantile values were 0% tipped. We see that the median values for Control and Treatment were 8.5% and 10.9% and the 3rd quantile values were 16.4% and 20.0% tipped, respectively.

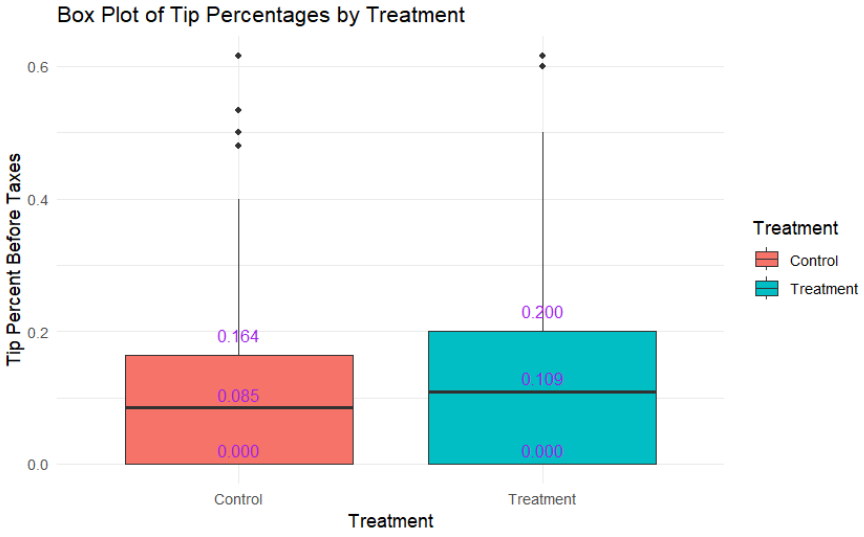

Figure 5. Box and Whisker Plot For Tip Percentages for Treatment vs Control

Day of Week	Average Tip ($)	Tip Standard Deviation ($)	Average Tip Percent (%)	Tip Percent Standard Deviation (%)	Number of Tippers	Percent of Transactions
Weekday	0.821	1.041	0.097	0.114	652	0.771
Weekend	0.937	1.148	0.093	0.106	191	0.229

Figure 6. Summary Statistics for Weekend vs Weekday

Instead, if we investigate tips on weekends versus weekdays, we see that Weekdays have an average tip of $0.82 and $0.94 on weekends. There are 350 customers on the weekends from the experiment and 1,181 customers on the weekdays. 191 and 652 of those weekend and weekday customers respectively tipped, which equates to 54.5% and 55.2% of customers tipping. Ultimately, through EDA, we see that Treatment days have more tip values and we also see that weekends yield higher tips, but we will investigate the significance of these increases in values.

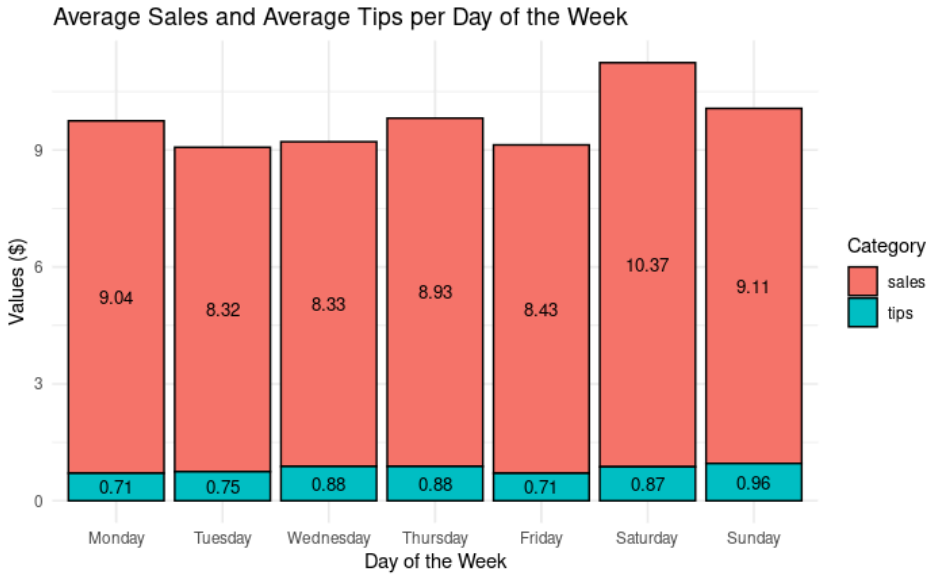

Figure 7. Average Amount of Sales and Average Tips by Day

We see in Figure 7. That there is on average, more spent per purchase on the weekends (Saturday and Sunday). We see that there is an equitable average amount of tips throughout the week, ranging between $0.71 and $0.96, where the highest tip average occurs on Sunday and the lowest average tip is tied for Monday and Friday.

Regression Analysis

To assess the effect of the appreciation sign on tipping behavior, we used a linear regression model. The dependent variable was the tipping percentage before tax, calculated as the total tips divided by the total sales for each transaction. The primary independent variable was the treatment indicator, which took the value of 1 for days when the appreciation sign was displayed and 0 for control days. The model controlled for other potential confounders such as transaction amount and day of the week.

$$\textbf{Tip Percentage}_i = 0.099 + 0.011 * \textbf{Treatment}_i + \epsilon_i$$

	Dependent variable:					
	Tip Percentage Before Tax					
	Baseline	Baseline Clustered SE	Inc. Transaction $	Inc. Transaction $ Clustered SE	Inc. DoW	Inc. DoW Clustered SE
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.011*	0.011	0.012**	0.012	0.020*	0.020
	(0.006)	(0.008)	(0.006)	(0.008)	(0.011)	(0.014)
Transaction Amount			−0.002***	−0.002***		
			(0.0004)	(0.0003)		
Weekday					0.011	0.011
					(0.010)	(0.016)
Treatment:Weekday					−0.012	−0.012
					(0.013)	(0.017)
Constant	0.099***	0.099***	0.112***	0.112***	0.090***	0.090***
	(0.004)	(0.007)	(0.006)	(0.008)	(0.008)	(0.014)
Observations	1,531	1,531	1,531	1,531	1,531	1,531
R^2	0.002	0.002	0.012	0.012	0.003	0.003
Adjusted R^2	0.002	0.002	0.011	0.011	0.001	0.001
Residual Std. Error	0.112 (df = 1529)	0.112 (df = 1529)	0.112 (df = 1528)	0.112 (df = 1528)	0.112 (df = 1527)	0.112 (df = 1527)
Note:						*p<0.1; **p<0.05; ***p<0.01

Figure 8. Regression Results

Average Treatment Effect:
The Average Treatment Effect (ATE) is 0.0108, meaning that on average, the tipping percentage on days with the appreciation sign was 1.08 percentage points higher than on control days. When utilizing robust standard errors, the regression provides a confidence interval between -0.04% and 2.2%. This effect, while positive, yielded a p-value of 0.0615, which is not statistically significant at the 5% level, but rather 10% (See column 1 of Figure 8). However, when we use clustered standard errors to account for our randomization method, the results are no longer significant (See column 2 of Figure 8).

Additionally, we theorized that people who purchase more expensive items or more goods may generally leave a higher tip due to the amount of work required for the barista. Therefore, we ran another regression in which we controlled for the price of the transaction, prior to the tip. As shown in column 3 of the regression table below, controlling for the price increases the ATE to be 0.012, or an increase in tipping percent of 1.2% on treatment days versus control which is also significant at the 5% level. However, using clustered standard errors changes the results to not be significant (See column 4 of Figure 8).

Heterogeneous Treatment Effects

To investigate whether the effect of the appreciation sign varies by day of the week, we extended our regression analysis to include interaction terms to indicate if the day is a weekday, or not.

$$\textbf{Tip Percentage}_i = 0.090 + 0.020 * \textbf{Treatment}_i + 0.011 * \textbf{Day of Week}_i$$

$$- 0.012 * (\text{Treatment}_i \times \text{Day of Week}_i) + \epsilon_i$$

This allows us to assess whether the treatment effect differs between these two types of days. The results of the regression can be found in column 5 (and 6 with clustered standard errors) of Figure 8 above.

In this case, the constant represents the baseline scenario, which is the average tip percentage on a weekend control day when the sign is not up, which has an average of a 9% tip. When compared to the baseline, the treatment increases tips by 2% which is significant at the 10% level. Additionally, when compared to baseline, weekday customers have an increase in average tips by about 1.1%. The interaction term indicates the scenario when it is a treatment day and a weekday, which actually decreases the average tip by about -1.2%, though not showing a significant result.

The heterogeneous treatment effects analysis reveals that the appreciation sign generally increases tipping percentages across all days of the week. There is a slight increase in the effect on weekdays, although not significant at a 5% level.

Post-Experiment Power Analysis

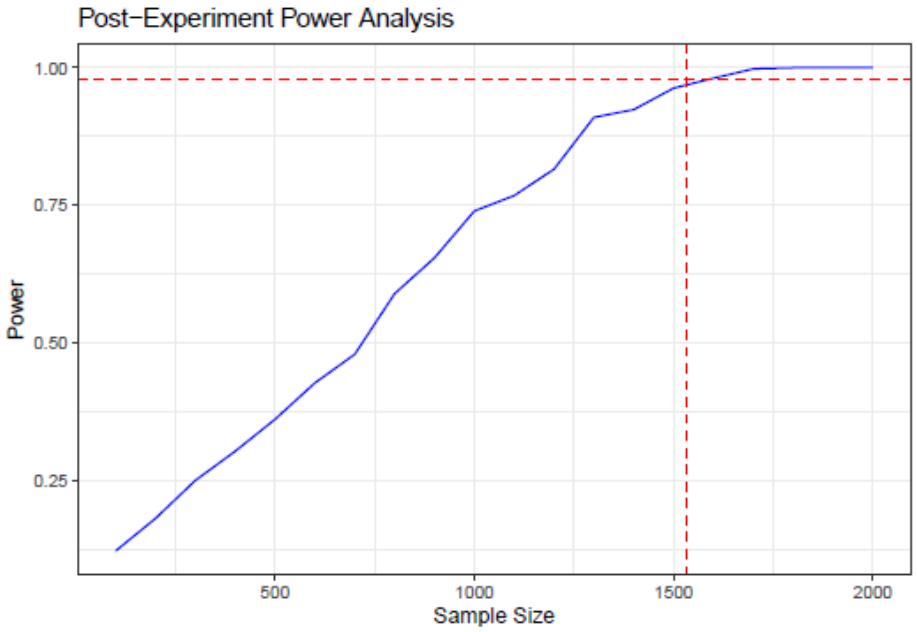

Figure 9. Post Experiment Power Analysis Graph

Because our results were not statistically significant, we ran a post-experiment power analysis to understand how many subjects we may have needed in order to determine a significant causal effect based on the collected data distribution. We determined our initial assumptions of the standard deviation of tips between the control and treatment group was far too low, around 3-5%,

when in fact our final collected data showed the standard deviation of each group being around 11%. The post-experiment power analysis suggests that with our 1,531 samples, we achieve about 85% power. While 85% power is often a strong result, we recommend increasing the sample size to above 2,000 due to the high variability in the data in order to achieve an even higher powered experiment. Additionally, we could also consider including more covariates, such as time of day or the barista to try to reduce the variability in tipping amounts.

Conclusion

After assessing the experiment we ran for two weeks which consisted of 1,531 purchases, we conclude that there is no statistical significance in the difference of tips between Treatment and Control groups at the 5% level, where the Treatment is defined as posting a sign of appreciation by the register that reads “Thanks for Supporting Local.” Additionally there is a positive effect on weekdays, though also not significant at the 5% level.

If we had more resources like time and money, we could run the experiment longer to harness more power in our analysis and to ensure an equal number of treatment and control transactions. While our tests did not show significance at at 5% level, offering words of gratitude to customers for supporting a local business remains a positive and appreciated practice.