Summer 5-15-2023

Essays on Customer Relationship Management

Nan Zhao
Washington University in St. Louis, judyzhao93@gmail.com

WASHINGTON UNIVERSITY IN ST. LOUIS

Olin Business School

Dissertation Examination Committee:
Raphael Thomadsen, Chair
Arun Gopalakrishnan
Xiang Hui
Chakravarthi Narasimhan
Song Yao

Essays on Customer Relationship Management

by
Nan Zhao

A dissertation presented to
Olin Business School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy in Business Administration

May 2023
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgments

I am deeply grateful to the people who have supported me throughout this journey. I would like to start by thanking my advisor Raphael Thomadsen for his mentorship. He has always been a role model and has taught me different ways to approach and solve problems. Next, I would like to thank my committee members and coauthors Chakravarthi Narasimhan and Arun Gopalakrishnan for their tremendous help and support while I was on the academic job market. I would also express my greatest gratitude to my committee member and coauthor Song Yao for being an amazing mentor on so many things. Moreover, I would like to thank my committee member Xiang Hui for always giving me helpful suggestions.

My gratitude also goes to my WashU friends who supported me during the job market: Zhenling Jiang, April Yu, Bicheng Yang, Yijun Chen, and Tianxin Zou. I am lucky to have them in my life and will forever remember what an awesome friend they have been.

Finally, I would never have become the person I am today without the unconditional love from my family and my husband. I look forward to the future we are going to spend together.

Nan Zhao

Washington University in Saint Louis

May 2023

Dedicated to my parents and my husband.

ABSTRACT OF THE DISSERTATION

Essays on Customer Relationship Management

by

Nan Zhao

Doctor of Philosophy in Business Administration

Washington University in St. Louis, 2023

Professor Raphael Thomadsen, Chair

The general topic of my dissertation is customer relationship management. Specifically, I use quasi-experimental causal inference methods, randomized field experiments, and machine learning methods to study and measure consumer response to co-branded credit cards and email communications promoting subscriptions.

In Chapter 1, “The Impact of Co-branded Credit Card Adoption on Customer Loyalty", we estimate the treatment effects of adopting a co-branded credit card on spending and loyalty behaviors using a comprehensive longitudinal dataset from a North American airline. Our data set contained detailed records of both airline credit card adopters and non-adopters, including their travel and loyalty program activities over a four-year horizon. We deal with the self-selection of card adopters by (1) using rolling-based matching procedure, (2) conducting difference-in-differences estimation on the matched sample with a two-way fixed effects specification, and (3) dividing treatment effects into three phases of time and argue that the endogenous timing of card adoption will most likely manifest in the short-term effect and is least likely to affect long-term effect. We find statistically significant and economically meaningful effects of card adoption on a multitude of behaviors. Specifically, flight spend was lifted by 42% when considering spend more than 12 months after adoption, demonstrating the persistence of the effect. These flight spend increases were largely driven by more flights

purchased rather than higher prices paid per flight, which is indicative of increasing share-of-wallet among adopters. Card adopters also increased award flight redemption to a greater extent than redeeming loyalty program points with airline partners. Finally, card adopters who experienced the highest increase in flight spend, tended to live near hub airports of the airline firm or were already existing members of the loyalty program.

In Chapter 2, “An Experimental Investigation of Price vs. Non-Price Messaging in Subscription Programs", we study how firms can attract and retain customers for subscription services. Subscriptions of digital and physical goods are becoming increasingly popular, and firms often compete heavily on price for customer acquisition. However, the challenge associated with advertised price discounts is substantial, as the featured price discounts highlight price savings and this might create an adverse selection problem with some customers signing up for just the price discount and then churning soon after. We worked with a major retailer that sells pet products, and we launched a four-week field experiment where we randomized price and non-price messaging in email advertising of subscription. We find that the non-price messages perform as well as the price messages in terms of sign-up rates and outperform price messages for reorder rates. This pattern also holds for number of orders, revenue and profit margin. We find that the inferior performance of the price message is primarily due to price attracting lower quality customers. Our findings suggest that one of the most dominant strategies of selling subscriptions is very suboptimal. Firms would be better off with the messages that make non-price motivations more prominent. Further, firms could also use previous purchase history to better target customers who could be a good match for the subscription services. Our results suggest that the price message should be sent to customers who are less familiar with the online channel, customers who are new to subscription, customers who have more regular purchase history, and customers who are less deal-prone. The rest should be sent non-price messages. Customers with no prior engagement with the firm should be targeted with the risk message by default. Finally, those who are most deal-prone and most familiar with online channels should not be sent any messages at all.

Chapter 1

The Impact of Co-branded Credit Card Adoption on Customer Loyalty

1.1 Introduction

Loyalty programs (LPs) are commonly used by firms with the intention of increasing customer engagement. Due to the proliferation of LPs, U.S. households collectively belonged to about 3.3 billion LP memberships in 2014 ([1]). While LP memberships continue to grow, less than half of the enrollees engage actively within these programs. As a result, firms are increasingly looking for new ways to enhance their LPs such as by partnering with other organizations. These partnerships typically take the form of joining a coalition LP where a set of firms jointly administer a program or offering a co-branded credit card (typically issued by a financial institution using a firm's brand on the card) to LP members. While a burgeoning literature (e.g., [2]; [3]; [4]; [5]) has documented the effects of coalition LPs on consumer behavior and loyalty, there is a dearth of work examining the impact of co-branded credit card adoption on new and existing LP members.

Co-branded credit cards are especially used by firms in the travel, hospitality, and retail industries to enhance customer engagement. For example, Macy's Star Rewards members can choose to adopt a co-branded American Express credit card and Marriott Bonvoy members have the option of signing up for a co-branded Chase Visa credit card. Similar examples abound in the airline industry with most airlines offering credit cards in partnership with financial institutions. Such credit cards offer customers the chance to earn LP points in three ways: (1) initial bonus points if spending requirements are met within a few months of card adoption, (2) in-network spend with the focal firm that can carry bonus multipliers on points earned, and (3) out-of-network spend with any other retailers that has a given exchange rate with the LP's points. In addition, exclusive perks may be offered to LP members who adopt the co-branded credit card. New customers, in addition to current LP members, could also sign up for such cards, which would automatically enroll them into the LP.

Industry reports suggest firms make healthy revenues by selling their LP points to credit card issuers who offer and manage the co-branded credit card.[1] However, there are still limited insights on how adoption of such a co-branded credit card influences subsequent customer behaviors with the brand itself. In addition, since consumers can earn an initial bonus with just out-of-network spend at other merchants, it remains unclear if becoming a card adopter has any impact on spend with the focal firm itself. The importance of this point cannot be understated. For example, suppose a co-branded credit card offers 20,000 initial bonus loyalty points if the card adopter spent $2,000 in the first three months. Nothing stops the adopter from putting their regular household expenses on this card during that timeframe

[1] https://getpocket.com/explore/item/airlines-make-more-money-selling-miles-than-seats

and earning the "free" initial bonus, without ever having to spend with the branded firm itself.

A number of questions relating to the impact of co-branded credit cards therefore remain unaddressed. First, and most crucially, firms contemplating the introduction of co-branded credit cards lack empirical evidence on their efficacy in lifting spend and loyalty with the brand. Second, whether any change in loyalty behaviors may persist over time is unclear. That is, would the effect of the card on adopters simply wear out after a period of time, such that any impact is short-lived? Third, what types of customers should be targeted by the firm? For example, would the firm be better off targeting already-loyal customers to further enhance their engagement, or use such a credit card to attract less-loyal customers? Fourth, understanding if card adopters make more frequent purchases or spend more on any given purchase would help managers position their messaging accordingly. Finally, how does engagement with other measures tracked by the firm's LP such as points earned and redeemed change with card adoption?

The purpose of this paper is to provide empirical evidence towards these questions and quantify the impact of card adoption, which is of high managerial relevance to firms looking to leverage co-branded credit cards.

We do so using a unique and comprehensive data set from a major North American airline (that wishes to remain anonymous) that operates its own loyalty program and also offers co-branded credit cards through partner financial institutions. Our single-source dataset is unusually rich in its descriptive features, containing records of 400,000 randomly sampled customers from the firm's LP over a four-year horizon including their flight bookings and LP usage. Importantly, our data set includes a subset of customers who adopted a co-branded credit card during our observational window, yielding both card adopters and non-adopters in our data.

The challenge in estimating treatment effects of credit card adoption comes from the self-selection of customers. As a result, directly comparing card adopters and non-adopters is likely to yield biased estimates because those who adopt the card may systematically differ from those who do not. For example, if card adopters tend to be customers who spent more with the firm in the previous year than non-adopters, post-adoption differences may reflect such selection differences rather than the causal impact of the card itself.

We adopt a multi-pronged approach to dealing with potential selection bias. First, we employ a rolling-based matching procedure in which non-adopters are matched to adopters using a battery of behavioral and demographic variables. Second, we conduct difference-in-differences (DID) estimation using the matched individuals' monthly spend with the airline (and also explore other behavioral measures). In DID estimation, we also control for individual and time fixed effects. As has been shown in other work on loyalty programs (e.g., [6]), a combination of matching and DID is helpful to mitigate selection effects. Our results should be viewed, of course, as the average treatment effect on the treated since they are conditional on credit card adoption.

We also explore possible selection effects of endogenous timing of card adoption, which may still persist after our rigorous controls above. An analogy to help explain the source of this endogeneity is to imagine a customer who waits to adopt a co-branded credit card, until just prior to a heavy schedule of pre-planned travel. Such an adopter can then make use of card perks (which can include free checked bags, priority boarding and upgrades, among others) as well as earn bonus multipliers on spend with the airline. In the extreme, should all adopters behave in this way, and significantly reduce travel on the airline after their pre-planned heavy schedule, the treatment effect from DID estimation may be upwardly biased and not reflective of the causal impact of card adoption.

Our approach to mitigating the potential endogenous timing of adoption is to divide treatment effects into three phases of time: less than three months after card adoption (which we label as the short-term), three months to less than twelve months after card adoption (which we label as the medium-term), and twelve months or more after card adoption (which we label as the long-term). We argue that endogenous timing of card adoption is most likely to affect short-term effects and least likely to affect long-term effects for the following reasons. First, most co-branded credit cards in the timeframe of our data carry an annual fee, which would dissuade customers from adopting too far in advance of any heavy travel period. Second, customers in our data set book flights 35 days prior to travel, on average (the 80% percentile is still only 55 days, which is shorter than our definition of short-term). This planning horizon is shorter than 70 to 76 days prior to travel, which is suggested by industry reports as the best time to book flights to secure a low price[2]. As a result, we argue that consumers are unlikely to be forward looking to the extent of timing card adoption around future flights that are more than 3 months out from adoption. Third, noting the difficulty of getting refunds from cancelling flights already booked, customers incur a considerable risk if they tie up their capital in flights far out in time that they may have to change plans on. For these reasons, we interpret treatment effects that persist beyond the short-term time phase as more compelling evidence of card adoption effects. We note that while card adoption effects may be present even in the short-term, the difficulty of teasing these apart from likely endogenous card adoption timing effects means that we focus on the medium- and long- term phases in terms of their implications for managerial practice.

[2]https://www.cheapair.com/blog/cheapair-coms-annual-airfare-study-reveals-the-best-time-to-buy-airline-tickets/

We also explore heterogeneous treatment effects of adoption in the medium- and long- term to better understand what types of customers show the highest lift from card adoption, which can help firms target the best prospects for co-branded credit cards. Dimensions we consider include customers living near hub airports of the focal airline versus non-hub airports, month of card adoption, and whether adopters were already members of the loyalty program or new-to-the-program.

We find statistically significant and economically meaningful effects from card adoption on monthly flight spend. In the medium and long term, we find that treatment effects on monthly flight spend persist with an average lift of 64% and 42% respectively. Monthly flight spend increases were driven largely by an increase in the number of flights purchased by adopters, rather than higher airfares per purchase. This finding is in line with a share-of-wallet expansion rather than upselling adopters to higher-value products.

Adopters also demonstrate higher engagement with the loyalty program, earning more miles, redeeming more award flights, and also redeeming more with partners of the airline (e.g., hotels, car rentals, other airlines). However, the effect size for award flights is higher than that for partner airlines. This finding suggests that adopting the credit card focuses more attention on using loyalty program miles on the focal airline as opposed to partners.

The highest revenue increase was realized from adopters who live in the metropolitan area where the focal airline has a hub airport (i.e., has the highest or second-highest market share of air traffic in that airport), or are existing LP members (as opposed to those who joined the LP along with the credit card).

Put together, these findings provide a novel perspective on the impact of co-branded credit cards that has not been previously discussed in the marketing or loyalty program literature.

Customers who adopted the card, in our study, were meaningfully more valuable than non-adopters over a sustained period of time and were more engaged with the loyalty program. Therefore, the impact of card adoption in our study went beyond the financial benefits to the firm of selling their loyalty points to banks that acquire such adopters.

The remainder of this paper is organized as follows. In Section 2 we connect our paper to the existing literature. In Section 3, we discuss institutional details and the descriptive features of our data set. In Section 4, we detail our matching procedure. In Section 5, we present our main findings. We conclude in Section 6.

1.2 Literature Review

Our work relates to and builds on three strands of literature on loyalty programs. The first strand is general work on loyalty programs and their effects. The second involves an emerging area of enhancing engagement in loyalty programs through partnerships. The third is specific to airline loyalty programs.

1.2.1 General Loyalty Program Literature

The loyalty program literature is mixed in terms of whether program members meaningfully change their behaviors. [7] finds that a loyalty program is more effective in getting light buyers to increase their purchases and that heavy buyers did not change their buying habits due to the program. [8] critiques the concept of loyalty programs as simply looking for short-term revenue gains from consumers rather than building long-lasting relationships while [9] do not find that loyalty programs create switching costs for consumers. [6] show that joining a loyalty program increases value by reducing customer attrition. [10] identify that loyalty program enrollees who generate the highest returns tend to live near a competitor's store such that they transfer spend from competitors to the store whose program they join. Further, a number of these studies point to the challenge of self-selection when analyzing data from loyalty programs since enrollment is typically not randomized. That is, a consumer who chooses to join a loyalty program may differ materially in terms of their future behaviors from one who does not.

Put together, it appears that joining a loyalty program may work for subsets of customers in preventing defection, in shifting spend from competitors, or by expanding the relationship with previously less established customers. However, since many firms have operated loyalty programs for decades, the managerial challenge is to keep or grow the engagement level with customers to fend off program stagnation. We next discuss the literature on enhancing engagement through partnerships in loyalty programs.

1.2.2 Enhancing Engagement in Loyalty Programs through Partnerships

Firms are increasingly looking to enhance the value of their loyalty programs to members by partnering with other organizations. One such partnership, of course, is a co-branded credit card which is the subject of our study, and has not been explored in the extant literature. A nascent literature has examined a different kind of partnership which involves a coalition loyalty program – in which multiple firms join a single program that customers can use to earn and redeem rewards. While such programs are more common outside of the United States (e.g., Flybuys in Australia, Air Miles in Canada), there are recent examples of such coalition LPs in the United States as well, such as Plenti, a multi-retailer program administered by American Express. We review this literature from the viewpoint of the potential benefit of partnerships in changing customer behavior.

[4] find that entry of a new firm into a coalition loyalty program generates spillover effects for customers at existing merchants and leads to higher sales. [2] show the importance of coalition loyalty program design for its profitability. [3] find that store affinity, in terms of geographic proximity or categories sold can generate positive financial effects in a coalition loyalty program. [5] examine a natural experiment in which a coalition loyalty program changed its reward earning structure (which made it more challenging to earn rewards) and find that the decline in spend is explained more by changes in the coalition network than customers changing their purchase patterns.

1.2.3 Airline Loyalty Programs

Our work also relates to and builds on the literature on airline loyalty programs. [11] and [12] suggest loyalty programs create switching costs and thus provide airlines more market power and higher profitability. [13] study the impact of participation in frequent flyer programs on customers' airline decisions. [14] and [15] find business travelers value the benefits of frequent flyer membership more than leisure travelers, and business travelers have a higher willingness to pay. [16] shows a frequent flyer program adds to an airline's competitive advantage and influences demand for fares on routes that are not dominated by the airline. Recently, [17] show that as travelers progress toward attaining tier status, they tend to fly more with the airline even when doing so costs more. They find the effect is more pronounced for business travelers who use business funds for expenses. Our work adds to the airline loyalty program literature by examining the impact of adopting a co-branded credit card on customer behavior, which has not previously been explored in this literature.

1.3 Institutional Details and Data Descriptives

In this section, we discuss the institutional details of the airline industry in the context of co-branded credit cards in Section 3.1. In Section 3.2, we provide an overview of the data set used in this study. In Section 3.3, we present a naive descriptive analysis (that does not control for selection effects) as a precursor to our main analysis.

1.3.1 Airline Industry and Co-branded Credit Cards

Airlines in North America are among the pioneers of loyalty programs which allow members to accrue a firm-specific currency (often called "miles") that can be used for redemption with the airline or partner retailers (e.g., hotels, rental cars, or other airlines). The earning of miles can also lead to tiered status by meeting requirements for spend, flight segments, or miles earned in a given calendar year. Earning higher tiers of status often allows LP members to access increasing perks such as priority boarding, free checked bags, priority upgrades, better seating, and concierge services.

The use of co-branded credit cards is prevalent in the airline industry. North American airlines, in particular, heavily promote such cards among potential adopters in a variety of media and at airports. In addition, travelers often discuss the pros and cons of co-branded credit cards on online forums such as FlyerTalk or read travel websites such as The Points Guy. As a result, a co-branded credit card is likely to have high awareness among potential adopters, which also highlights the need to carefully handle selection effects that can ensue. The research questions we aim to study are therefore highly relevant in this industry, which is part of the travel and hospitality sector of the economy. The potential benefits to customers signing up for such cards are multifold. Although these cards typically carry an annual fee, they offer the potential to earn an initial bonus of loyalty points (e.g., 30,000 to 50,000 miles), which are often equivalent to a round-trip ticket whose value alone exceeds the annual fee. To receive the initial bonus, customers typically need to meet a spending requirement (through in- or out-of-network spend) within a few months of adoption. In addition, card members may also qualify for travel perks such as priority boarding, free checked bags, and upgrades. Card holders earn points on every dollar spent using the card, often with multiple (bonus) miles per dollar spent with the airline.

Airlines also receive multiple benefits by offering co-branded credit cards. For example, airlines can sell miles to their bank partners, from which customers get airline miles back on each transaction (instead of cash). Major U.S. airlines earn billions in yearly revenues from selling miles to bank partners.[3] Moreover, airlines can also get sign-up bounties for each new co-branded credit card holder. Banks can also offer rebates on processing fees to airlines on qualified transactions, which provides further incentives for an airline to partner with them on credit cards.[4] Further, airlines can be strategic about restricting access to mileage redemption to specific flights and times in order to optimize their revenues.

1.3.2 Data Overview

Our data comes from a major North American airline that chooses to remain anonymous. The dataset contains detailed records of a random sample of 400,000 U.S.-based LP members in the timeframe of Jan 2015 to Dec 2018. Of these, 297,866 members did not adopt a co-branded credit card throughout the timeframe of our dataset. The remaining members (about 25% of the sample) either already held or adopted a card during our data window.

[3]Source: https://www.forbes.com/sites/advisor/2020/07/15/how-airlines-make-billions-from-monetizing-frequent-flyer-programs/?sh=5b3d4a9614b2

[4]Source: https://www.fool.com/investing/general/2015/04/18/heres-why-airline-co-branded-credit-cards-are-so-v.aspx

To adequately analyze treatment effects of adoption, we imposed a minimum of 12 months of data before and after adoption, to obtain a subset of 17,156 card adopters who adopted in 2016 or 2017 (7,895 adopted in 2016 and 9,261 adopted in 2017). The remaining card adopters either already adopted prior to 2015, in 2015, or in 2018, which would preclude a sufficient period of time both before and after adoption to carefully tease apart treatment effects from other confounding effects. Going forward, we present the summary statistics and analyses using the 315,022 LP members (17,156 adopters plus 297,866 non-adopters) that are the focus of this study.

For each member, we have data on (i) their flight bookings during the sample window, including date of travel, origin and destination, price paid, and miles earned; (ii) award redemptions including date, number of miles used, and nature of redemption (flight, upgrade, partner airlines, other partners and others); (iii) the miles earned through credit card adoption and usage, as well as other non-flight activities; and (iv) demographics including program joining date, gender, zipcode, yearly tier status, email subscription opt-in, and whether a member is classified by the airline as a business or a leisure traveler (based on the airline's proprietary algorithm). Importantly, because our data comes from the airline and not the credit card company, we have flight bookings for each member regardless of the method of payment.

In Table 1.1, we provide summary statistics on members' activities by segments from 01/01/2015 to 12/31/2018. These activities are summed over a four-year period and then averaged across members in each segment. Means and standard deviations (which are in parentheses) are reported in each column. We observe considerable differences between card adopters and non-adopters in this period. Specifically, card adopters take nearly double the number of paid flights, triple the number of award flights, and their earned miles and flight spend are also nearly double that of non-adopters. These differences may arise due to a

Table 1.1: Summary Statistics over 2015-2018

	All Customers	Adopters	Non-adopters
Number of Customers	315,022	17,156	297,866
Number of Flights	8.01	14.65	7.62
	(17.67)	(29.71)	(16.63)
Number of Paid Flights	7.43	13.07	7.10
	(16.88)	(28.47)	(15.90)
Number of Award Flights	0.58	1.58	0.52
	(1.95)	(2.83)	(1.87)
Number of Partner Redemptions	0.16	0.28	0.15
	(0.69)	(1.08)	(0.66)
Miles Earned thru Flights	12,119.05	23,739.68	11,449.75
	(53,456.97)	(85,980.37)	(50,874.82)
Flight Spend ($)	1,970.95	3,427.45	1,887.07
	(6,335.96)	(9,763.34)	(6,069.38)

Note: Means and standard deviations (in parentheses) are reported in each column

number of reasons including fundamental differences between adopters and non-adopters, the treatment effect of adopting a credit card, and temporal effects.

1.3.3 Naïve Descriptive Analysis

In this section, we present a naïve descriptive analysis in which we use card adopters and non-adopters in a difference-in-differences (DID) regression. We use monthly flight spend as the dependent variable as an example. We use a two-way fixed effects specification that controls for individual and time (at monthly level) fixed effects, and estimate the Equation 1.1 below:

$$FlightSpend_{it} = \gamma_i + \mu_t + \lambda \cdot Adoption_{it} + e_{it}, \tag{1.1}$$

where γ_i and μ_t are individual and year-month fixed effects, and $Adoption_{it}$ is the monthly credit card adoption status, which equals one for all credit card holders after signing up for the card and zero prior to sign-up, and is zero for all non card holders across all 48 months.

Table 1.2: Diff-in-diff Estimation Results

	Dependent variable:
	Flight Spend
Adoption	30.723***
	(1.381)
Observations	15,121,056
R^2	0.262
Individual FE	Y
Year-month FE	Y
Note:Standard errors are clustered at the individual level	*p<0.05; **p<0.01; ***p<0.001

We present the DID results in Table 1.2, from which it appears that adopters seem to spend $31 per month more than non-adopters on flights with the focal airline, after they sign up for the credit card. On an annual basis, this amounts to over $360 in additional spend, which is equivalent to about an 80% lift as compared to non-adopters.

However, it would be naïve to conclude that all of this lift accrues from card adoption as there can be sources of selection bias beyond controls for individual and time fixed effects. We divide these sources into two buckets. The first involves observable differences between adopters and non-adopters. For example, adopters perhaps may have flown more frequently with the focal airline in the past, or may be from certain demographics (e.g., living close to a hub airport) that travel more. Failing to account for these differences can lead to inflated treatment effects of card adoption. We discuss a matching strategy to control for such differences (both time-invariant and time-varying ones) in Section 4. In this approach, we estimate a logistic regression for the probability of adoption as a function of a battery of covariates and obtain a set of adopters and non-adopters with closely matching distributions of observable covariates just prior to adoption.

The second source involves endogeneity in the timing of card adoption because the decision of when to adopt (i.e., $Adoption_{it}$ in equation 1.1) may be correlated with future travel on the airline. For example, suppose a customer knows that a high-intensity travel schedule on the focal airline is coming up in near future. It may then be valuable for this customer to adopt the card prior to this travel schedule in order to maximize the value of its benefits and perks. As a result, the random error terms of the adoption decision and spend conditional on adoption may be correlated even after matching. In Section 5, we present our empirical strategy and results in dividing treatment effects into short-, medium-, and long- term effects. We argue that endogeneous timing is most likely to affect short-term effects given the planning horizons observed in our data set (flights booked about 35 days prior to travel) and the incentives and fees associated with the credit card. Further, endogenous timing effects would also be indicated if the effects are short-lived and do not persist over a longer time horizon after card adoption.

1.4 Matching of Card Adopters and Non-adopters

Instead of using all adopters and non-adopters in our data set for analysis, we sought to obtain a control group with a closely matching covariate distribution as compared to the adopter group ([18]). The literature suggests two popular approaches for matching each adopter with a suitable non-adopter: propensity score matching (e.g., [19]) where a scalar measure is obtained to compare adopters and non-adopters, and coarsened exact matching (e.g., [20]) where variables are converted to categorical levels with the aim of seeking an exact match between an adopter and a control non-adopter. We use propensity score matching as we include a large number of behavioral variables that are essential to adoption, which becomes challenging to deal with in exact matching (which is more suitable in lower-dimensional covariate spaces).

In this approach, we use the group of customers who adopted the airline's credit card in 2016 and 2017 as the treatment group. We match customers who adopted the card at month t ($01/2016 \leq t \leq 12/2017$) with customers who had very similar probabilities of adoption at t, i.e., essentially the same hazard rate, but who did not end up signing up for the card during the observed period based on a propensity score constructed from a set of demographic and behavioral variables.[5] We summarize these matching variables in Table 1.3.

Our eighteen matching variables consist of both demographics and behavioral activities describing the trajectory of each customer up to month t. Of these, gender,[6] email opt-in status (those subscribed to the airline's emails may be more likely receive credit card offers), business (business travelers may have higher demand for air travel), hub (those living in the metropolitan area of a hub airport have more opportunities to travel with the focal airline), tier status (members with and without status enjoy different benefits with the focal airline), and mileage balance (members who already possess a large number of miles might have different incentives than those with fewer miles) highlight customer demographics that are most likely to be imbalanced between adopters and non-adopters. The remaining variables in Table 1.3 capture the engagement level and consumption intensity of each member up to month t. Specifically, we distinguish activities and consumption in the last 3 months (i.e., $t-1$ to $t-3$) from those in the last 4-12 months (i.e., $t-4$ to $t-12$) to discriminate between recent behavioral patterns that might differ from those earlier in time, and to account for the possibility that adopters could become even more engaged as they were close to adopting the airline's credit card.

[5] Our primary approach of estimating the probability of adoption is a logistic regression. We also used a probit regression for robustness, and the results are very similar.

[6] Gender differences have been found in credit limit and credit card usage. For example, https://www.paymentsjournal.com/men-and-women-and-credit-cards/, and [21]

Table 1.3: Descriptions of Matching Variables

Matching Variable	Variable Type	Description
Gender	Categorical	Female, Male or Unknown
Email Opt-in	Binary	Whether a customer opted in for emails
Business	Binary	Whether a customer was classified by the airline as a business or leisure traveler
Hub	Binary	Whether a customer's zipcode belongs to the metropolitan area of each hub airport
Tier Status	Binary	The most recent tier status (statused or not) of each customer prior to adoption month t (for adopters), or a randomly assigned month that matches t (for non-adopters)
Mileage Balance	Binary	The most recent mileage balance ($\geq 50,000$ or not) of each customer prior to adoption month t (for adopters), or a randomly assigned month that matches t (for non-adopters)
Last 3 Months' Flights Taken	Numeric	Total number of flights taken in the last 3 months (paid and award) of each customer prior to adoption month t (for adopters), or a randomly assigned month that matches t (for non-adopters)
Last 3 Months' Flight Spend	Numeric	Total flight spend ($) in the last 3 months of each customer prior to adoption month t (for adopters), or a randomly assigned month that matches t (for non-adopters)
Last 3 Months' Award Flights Taken	Numeric	Total number of award flights taken in the last 3 months of each customer prior to adoption month t (for adopters), or a randomly assigned month that matches t (for non-adopters)
Last 3 Months' Redemptions	Numeric	Total number of redemptions made in the last 3 months (flights and partners) of each customer prior to adoption month t (for adopters), or a randomly assigned month that matches t (for non-adopters)
Last 3 Months' Miles Purchased	Numeric	Total number (frequency) of airline miles purchased in the last 3 months of each customer prior to adoption month t (for adopters), or a randomly assigned month that matches t (for non-adopters)
Last 3 Months' Miles Earned thru Flights	Numeric	Total number of airline miles earned thru flights in the last 3 months of each customer prior to adoption month t (for adopters), or a randomly assigned month that matches t (for non-adopters)
Last 4-12 Months' Flights Taken	Numeric	Total number of flights taken in the last 4-12 months (paid and award) of each customer prior to adoption month t (for adopters), or a randomly assigned month that matches t (for non-adopters)
Last 4-12 Months' Flight Spend	Numeric	Total flight spend ($) in the last 4-12 months of each customer prior to adoption month t (for adopters), or a randomly assigned month that matches t (for non-adopters)
Last 4-12 Months' Award Flights Taken	Numeric	Total number of award flights taken in the last 3 months of each customer prior to adoption month t (for adopters), or a randomly assigned month that matches t (for non-adopters)
Last 4-12 Months' Redemptions	Numeric	Total number of redemptions made in the last 4-12 months (flights and partners) of each customer prior to adoption month t (for adopters), or a randomly assigned month that matches t (for non-adopters)
Last 4-12 Months' Miles Purchased	Numeric	Total number of airline miles purchased in the last 4-12 months of each customer prior to adoption month t (for adopters), or a randomly assigned month that matches t (for non-adopters)
Last 4-12 Months' Miles Earned thru Flights	Numeric	Total number of airline miles earned thru flights in the last 4-12 months of each customer prior to adoption month t (for adopters), or a randomly assigned month that matches t (for non-adopters)

Since adopters could have adopted the card at any month between 01/2016 and 12/2017, and we observe credit card adopters in all 24 months, matched pairs must be formed such that adopters and non-adopters are similar *just prior to* the month of adoption. Therefore, instead of using a common initialization period (e.g., [22]) for all customers, we use the approach of rolling entry matching (e.g., [23]). That is, we calculate the behavioral activities and consumption intensities prior to the exact month of adoption - month t for all 17,156 adopters. Further, we randomly divide all 297,866 non-adopters into 24 groups, with each group randomly assigned one month t out of the 24 months, such that we could use the assigned month t to calculate the prior activities for all non-adopters, respectively. In other words, because non-adopters don't have an actual adoption date, the month t assigned to each non-adopter sub-group serves to resolve the "missing treatment date" issue for control members. There are two benefits to our approach. First, we can get more accurate behavioral measures of adopters prior to adoption than just using a common initialization period (e.g., 01/2015-12/2015) since much could have changed between the end of the initialization period and the adoption month. Second, by randomly dividing non-adopters into 24 groups, we avoid re-sampling the same non-adopters for matching, which is a luxury we are able to incorporate due to the large size of the non-adopter sample.[7]

After preparing the relevant matching variables, we execute the propensity score in two steps. First, we compute the credit card adoption propensity for each monthly group of adopters and non-adopters as a function of the set of variables described in Table 1.3. Specifically, we estimate Equation 1.2 below:

$$Pr(Adoption_{it} = 1) = Pr(\omega + \delta X_{it} + e_{it} > 0), \tag{1.2}$$

[7] For robustness, we also used risk-set matching (e.g., [24]), where one customer can serve as control at different times for more than one treated customer and obtain very similar results.

where X_{it} is a vector of customer-specific characteristics, which include demographics and behavioral patterns up to month t. We assume e_{it} is distributed *i.i.d.* type I extreme value, leading to a binary logit model for the propensity score. We then fitted a separate logistic regression for each group of adopters and non-adopters who have the same assigned month t.

Second, we used the nearest neighbor matching algorithm, in which each adopter was matched with the closest non-adopter in terms of propensity score using the pair-wise Mahalanobis distance metric (with a caliper of 0.001, which is the maximum allowed distance between two matched customers).

From the initial sample of 17,156 adopters and 297,866 non-adopters, our matching procedure yielded 15,769 treated and 15,769 control customers, respectively. To check the quality of matching, we visually demonstrate covariate balance in standardized mean differences in Figure 1.1.[8]

From Figure 1.1, we can see that there were sizable differences in the observed dimensions between adopters and non-adopters before matching. After matching, the two groups are much more balanced and all covariates are within a standardized mean difference of 0.05.[9]

We also visually depict the covariate distributions before and after matching between adopters and non-adopters in Appendix A.1. All covariates are balanced and highly overlapping between adopters and non-adopters after matching.

In Figure 1.2, we provide a visual description of our goal of estimating how card adoption (the treatment) affects future behavior. Suppose a treated customer (shown in the first row) adopts in month t, matching ensures a control customer (shown in the second row) with a

[8]Standardized mean difference (SMD) is the difference in means of each variable between adopters and non-adopters divided by the pooled standard deviation.

[9]An SMD smaller than 0.1 is widely accepted to indicate a negligible difference in covariate balance (e.g., [25].

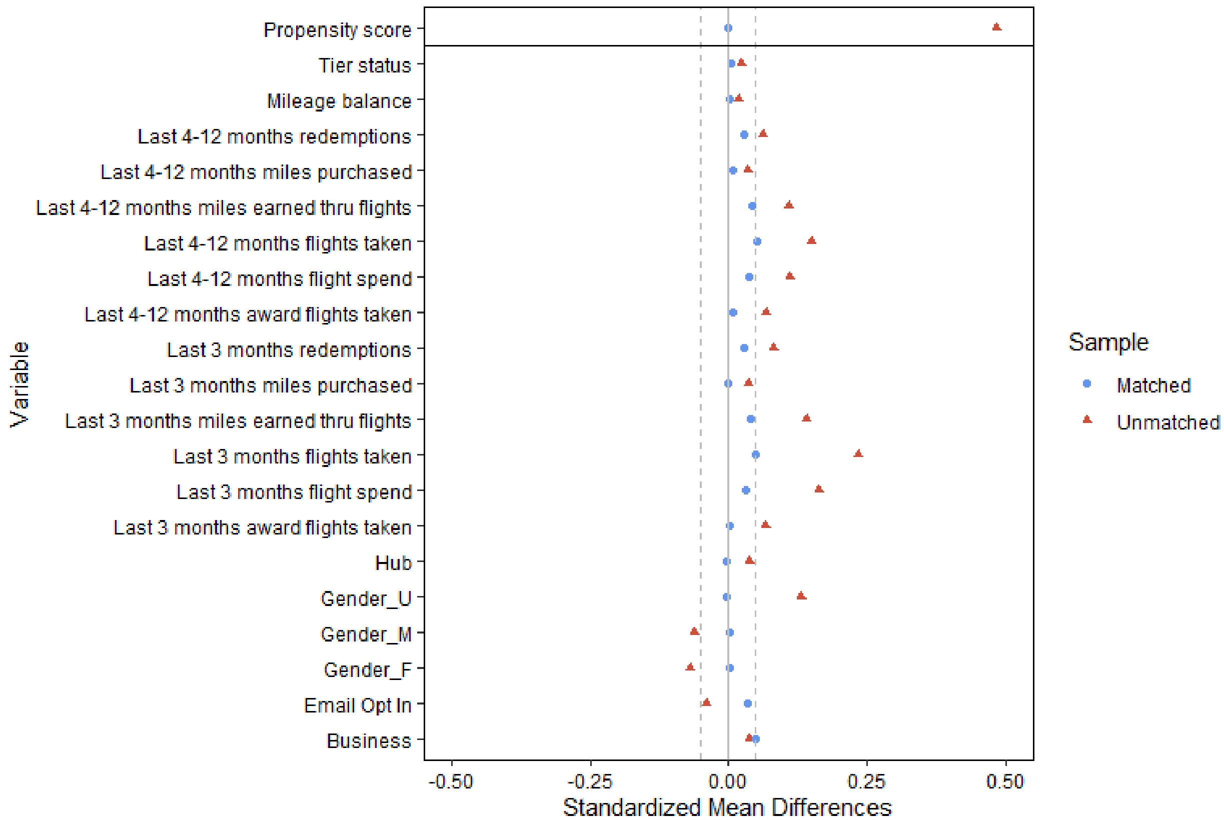

Figure 1.1: Comparison of Adopters and Non-adopters Between Matched and Unmatched Sample

very similar previous behavioral history and demographics. Differences that ensue after the month of adoption between the treated and control customer (over and above individual and time fixed effects) can then identify the treatment effect if the observables used for matching resolve all selection concerns. The question mark denotes the need to identify the size and direction of any treatment effect. Of course, should other sources of selection bias be present (e.g., endogeneous timing of adoption based on unobservables), this would need to be accounted for as well, which we discuss in Section 5.1.2.

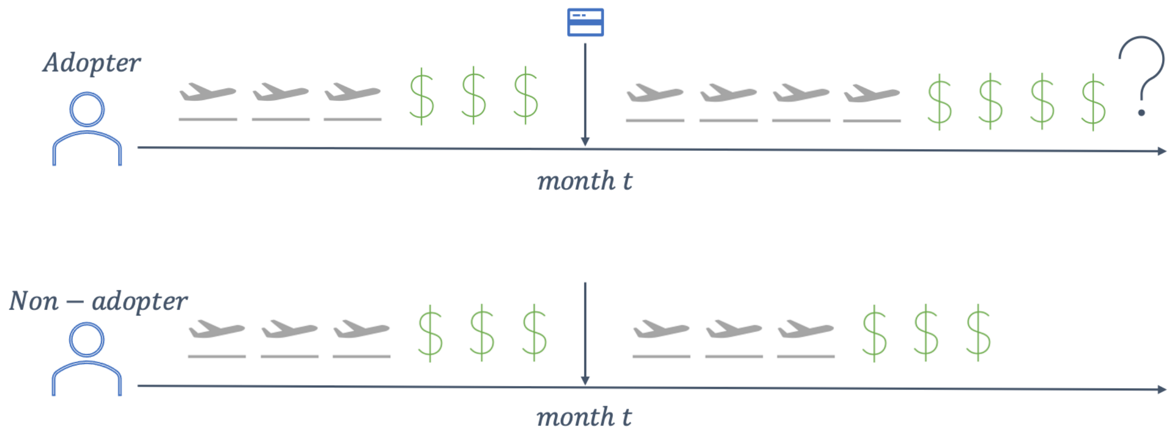

Figure 1.2: Matching and Possible Future Trajectory

1.5 Empirical Strategy and Results

We begin by describing our empirical strategy in Section 5.1. In Section 5.2, we present our average treatment on treated effects. In Section 5.3, we discuss heterogeneous treatment effects. We present robustness checks in Section 5.4.

1.5.1 Empirical Strategy

Our empirical strategy is to estimate causal effects of card adoption on a variety of dependent variables relating to customer behavior with the focal airline. In Section 5.1.1 we demonstrate that the matching procedure described in Section 4 yields null results in a placebo test ([18]), providing reassurance of adherence to the parallel trends assumption in DID estimation. Next, in Section 5.1.2, we discuss how we deal with endogenous timing of card adoption that may upwardly bias estimates of card adoption effects. Our treatment effects should be interpreted, of course, as the average treatment effect on the treated (ATT) (e.g., [26], [27]).

Placebo test for parallel trend

Table 1.4: Placebo Tests for Parallel Trend After Matching

	Dependent variable:				
	Num of Paid Flights	Num of Award Flights	Miles earned thru Flights	Flight Spend	Num of Partner Redemptions
	(1)	(2)	(3)	(4)	(5)
Placebo Adoption	0.005	0.0004	12.703	1.220	0.0003
	(0.003)	(0.001)	(7.577)	(1.082)	(0.0003)
Observations	378,456	378,456	378,456	378,456	378,456
R^2	0.321	0.122	0.346	0.291	0.116
Individual FE	Y	Y	Y	Y	Y
Year-month FE	Y	Y	Y	Y	Y

Note:Standard errors are clustered at the individual level *p<0.05; **p<0.01; ***p<0.001

Unlike the naïve DID regression shown in Table 1.2, we use the matched sample obtained in Section 4.1 to account for selection on observables and to increase the chances of adherence to the parallel trends assumption of DID estimation.

We conduct a placebo test as a robustness check of whether parallel trends appears to be a reasonable assumption in our setting. Specifically, we take the 12 months of data prior to adoption (at time t) and insert a placebo adoption in the middle (i.e., at the 6 month mark). Should matching be effective, this placebo adoption, which equals one for adopters from $t-6$ to $t-1$ and zero otherwise, and is zero for non-adopters, should have no effect on the variables of interest. On the other hand, it would be concerning if such a placebo test yields a treatment effect indicating a divergence of treated and control groups in the pre-adoption phase itself.

Examining number of paid flights, number of award flights, miles earned thru flights, flight spend ($), and number of partner redemptions (e.g., hotels, car rentals, etc.) in Table 1.4, all at the monthly level, we find no effects for the placebo adoption in the pre-adoption period, which is reassuring.[10]

Endogeneous timing of card adoption

[10]We also report placebo test results on the unmatched sample in Appendix A.2, where parallel trend is violated for majority of the DVs.

Because co-branded credit card adopters may time card adoption to coincide with upcoming heavier-than-usual travel, a selection bias can result due to positive correlation in the error terms of the adoption decision and post-adoption travel behaviors. This bias, because it arises from unobservables, would not be controlled for by the variables used for matching (as discussed in Section 4). Our approach to dealing with this endogenous timing of card adoption is to consider the plausible planning horizon of consumers in terms of airline travel.

In our data set, flights are booked about 35 days (on average) prior to travel. Industry reports suggest that the best flight prices are obtained about 70 to 76 days prior to travel in the timeframe of our data set (see footnote 2). In terms of credit card applications, approvals are often made very quickly (within 1 to 2 days), and the card itself is typically mailed out to a new adopter within 7 to 10 business days.[11] Put together, a consumer who wishes to take advantage of card adoption for pre-planned travel can readily do so within a three-month (or 90-day) planning horizon. We therefore assume that post-adoption effects in the first three months are most prone to this form of selection bias.

Given the discussion above we assume that effects more than 3 months after adoption are less susceptible to this endogeneous timing selection effect for the following reasons. First, flight planning and bookings more than three months out inherently deal with uncertainty. Consumers may change their travel plans or they may fall through. In the timeframe of our data set, changes to flight bookings often came with substantial fees, which is a disincentive to plan travel too far in advance. Second, the initial bonus is awarded based on spending in the first few months, giving consumers an incentive to spend more soon after card adoption. Further, co-branded credit cards come with an annual fee which is typically charged in the first billing cycle after adoption. If any pre-planned travel is further out in time, it would be

[11] https://smartasset.com/credit-cards/how-long-does-it-take-to-get-a-credit-card

more likely that a rational consumer would prefer to wait till closer to such travel than to incur an up-front cost earlier and wait for a long time to reap the benefits.

We therefore divide the treatment effect into short-term (1-3 months after adoption), medium-term (4 to 12 months after adoption), and long-term (13 months or more) phases. As such, by examining the persistence of the treatment effect in the medium and especially in the long term, we obtain effects that are more plausibly and credibly attributable to card adoption. The pattern of effects going from the short- to medium- to long- term will also be revealing. In the event of a large short-term effect and zero treatment effects in the medium- and long-term, it could be argued that endogenous timing of card adoption largely drives post-adoption behaviors. On the other hand, the persistence of treatment effects in the medium- and long- term would point towards card adoption having an impact on behavior, even while the short-term may contain substantial selection bias. Our approach is therefore to focus on the medium- and long- term treatment effects in our main analysis.

For each dimension of customer behavior we measure the short-, medium-, and long-term effects, as defined by Equation 1.3 below:

$$\begin{aligned} Y_{it} = \alpha_i + \theta_t + \beta_1 \cdot I(1 \leq Months\_Since\_Adoption_{it} \leq 3) \\ + \beta_2 \cdot I(4 \leq Months\_Since\_Adoption_{it} \leq 12) \\ + \beta_3 \cdot I(13 \leq Months\_Since\_Adoption_{it}) + e_{it}, \end{aligned} \tag{1.3}$$

where Y_{it} is denotes the outcome of interest, such as flight spend, number of paid (purchased) flights, spend per flight, number of award flights, miles earned through flights, and number of partner redemptions, all measured at the monthly level. α_i and θ_t are individual and year-month fixed effects, and $Adoption_{it}$ is the monthly credit card adoption status, which

Table 1.5: Diff-in-diff Estimation Results on the Matched Sample

	Dependent variable:
	Flight Spend
Adoption	27.291***
	(1.140)
Observations	1,513,824
R^2	0.200
Individual FE	Y
Year-month FE	Y
Note:Standard errors are clustered at the individual level	*p<0.05; **p<0.01; ***p<0.001

equals one for all credit card adopters after signing up for the card and zero prior to sign-up, and is zero for all non card holders across all 48 months.

1.5.2 Average Treatment Effects on the Treated

We begin by presenting results for the key dependent variable of monthly flight spend (prior to dividing into three time phases). All of our analyses use a two-way fixed effects specification that controls for time-invariant customer characteristics, as well as common time trends and month-to-month fluctuations. Table 1.5 shows the result for monthly flight spend as the dependent variable. By directly comparing the estimates in Table 1.2 and Table 1.5, we observe that the point estimate is about 12% smaller after controlling for selection on observables. The p-value of the difference between the estimates of Table 1.2 and 1.5 is 0.055 (computed as per [28]).

Dividing treatments into temporal phases

Table 1.6: Diff-in-diff Estimation Results on the Matched Sample with Three Temporal Phases

	Dependent variable:
	Flight Spend
Adoption*short	52.274***
	(1.494)
Adoption*medium	23.957***
	(1.274)
Adoption*long	16.386***
	(1.392)
Observations	1,513,824
R^2	0.201
Individual FE	Y
Year-month FE	Y
Note:Standard errors are clustered at the individual level	*p<0.05; **p<0.01; ***p<0.001

We now break down this effect into temporal phases - short, medium, and long term in Table 1.6. The monthly spend effects are, respectively, \$52, \$24, and \$16 in the three phases of time.

Our first observation is that treatment effects persist in the medium- and long- term and are economically meaningful. Over the 9 months of the medium-term, an adopter spends \$216 more than a non-adopter on flights with the focal airline. This represents a substantial lift of 64%. To be more conservative, a lift of 42% still persists in the long term. The long-term revenue increase on an annualized basis is \$197.

Second, while attenuation in effect sizes over phases of time is not unexpected, that the long-term effect size is about two-thirds of the medium-term effect size is noteworthy in that card adoption effects have not largely dissipated even a year or more after adoption.

Third, the pattern of effect sizes in Table 1.6, we argue, are not entirely explained by endogenous timing of the adoption decision. Given the planning horizon of within three months that we discussed in Section 5.1.2, an effect size purely driven by endogenous timing should have resulted in a large short-term effect (which we find) but also close-to-zero effects beyond the short-term. The short-term effect size of $52 is likely a combination of endogenous timing and actual card adoption effects, but we are unable to tease apart their relative contribution, and focus our attention on the medium- and long- term effects in terms of their managerial relevance.

For reassurance, we further divide long-term (13 months +) into 13-24 months and 25-36 months post-adoption. We present our results in Table 1.7 below. The rationale of doing so is that forward-looking to the extent of more than 2 years is extremely difficult. We can see that the effect size plateaued at $16 per month after 13 months post-adoption, supporting our argument that the endogenous adoption timing is least likely to affect long-term results. Thus, we are more confident to conclude that $16 per month lift in flight spend is the true treatment effect of card adoption.

Table 1.7: Diff-in-diff Estimation Results on the Matched Sample with Four Temporal Phases

	Dependent variable:
	Flight Spend
Adoption*(1-3)m	52.277***
	(1.492)
Adoption*(4-12)m	23.966***
	(1.267)
Adoption*(13-24)m	16.339***
	(1.397)
Adoption*(25-36)m	16.647***
	(2.034)
Observations	1,513,824
R^2	0.201
Individual FE	Y
Year-month FE	Y
Note:Standard errors are clustered at the individual level	*p<0.05; **p<0.01; ***p<0.001

Other dependent variables

We now examine the ATT for other dependent variables including how changes in flight spend may be decomposed into number of flights purchased per month, and spend per flight. This decomposition helps inform whether card adoption effects lie in the realm of increasing purchase frequency with the airline or in increasing the amount spent on each flight. [17] show that both of these drivers can be in play for airline travelers who are looking to achieve elite status in their airline loyalty program. It is therefore of interest to examine how adoption of an airline credit card may impact these drivers. We also examine effects on loyalty program behaviors such as miles earned, redemption of award flights (using miles), and number of partner redemptions. These results are shown in Table 1.8, again divided by the three phases of time.

The ATT of 0.095 flights per month in the medium term translates to about 0.86 more flights over a 9-month time frame. In the long term, adopters take 0.77 more flights over a 12-month time frame. Taking the average price per flight of $258 in our dataset, we find that practically all of the effects on flight spend that we find in Section 5.2.1 are due to an increase in the number of flights purchased. It is therefore not surprising that spend per flight is largely unaffected by card adoption. While a small (and statistically significant) effect is observed in the medium term of $9.77, this is a change of less than 4% when compared to the average price per flight in our data set. Put together, card adoption appears to increase the flights purchased with the focal airline, which then drives increased spend.

We should expect that adopters having purchased more flights, should also earn more LP miles from these flights. However, it is important to note that miles can also be earned through credit card spend on non-airline purchases for adopters. Because non-adopters do not have access to this latter type of mileage earning, we only include airline miles earned

Table 1.8: Diff-in-diff Estimation Results on the Matched Sample with Temporal Phases

	Dependent variable:				
	Num of Paid Flights	Spend per Flight	Miles Earned thru Flights	Num of Award Flights	Num of Partner Redemptions
	(1)	(2)	(3)	(4)	(5)
Adoption*short	0.239***	−1.383	290.693***	0.013***	0.003***
	(0.005)	(3.833)	(10.065)	(0.001)	(0.0004)
Adoption*medium	0.095***	9.769*	156.152***	0.039***	0.006***
	(0.004)	(4.551)	(9.875)	(0.001)	(0.0004)
Adoption*long	0.064***	0.497	117.770***	0.040***	0.006***
	(0.004)	(5.114)	(11.675)	(0.001)	(0.0004)
Observations	1,513,824	163,182	1,513,824	1,513,824	1,513,824
R^2	0.251	0.336	0.238	0.058	0.052
Individual FE	Y	Y	Y	Y	Y
Year-month FE	Y	Y	Y	Y	Y

Note:Standard errors are clustered at the individual level *p<0.05; **p<0.01; ***p<0.001

through flights to facilitate a fair comparison. From column 3 of Table 1.8, we indeed find this to be the case: card adopters earn 156 more miles per month in the medium-term, and 118 more miles per month in the long-term.

Unlike flights purchased, redemption of LP miles for award travel is likely to reflect higher redemption over time (rather than a decay) since consumers may be more likely to redeem their miles for flights as their stock of LP miles increases. In column 4 we find this increasing pattern chronologically over the three phases of time. Interestingly, award flights redeemed per month plateaus after the medium-term.

Finally, the airline's LP miles can also be redeemed through partners such as other airlines, hotels, rental cars, and retail firms. Such partner redemptions may be substitutes to award flights with the focal airline since the same stock of miles can be used for either category of redemption. Similar to award flights, we find an increasing pattern of partner redemption (which is explained by the greater stock of miles built over time). However, the effect saturates between the medium- and long- term. The magnitudes in columns 4 and 5 suggest that adopters focus more of their mileage redemption activity on award flights with the focal airline as opposed to partners.

Table 1.9: Flight Spend

	Segment:										
	All	Hub	Non-hub	High Mileage	Low Mileage	High Bonus	Low Bonus	Peak Season	Off-peak Season	New Members	Existing Members
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Adoption*short	52.274***	55.446***	50.872***	54.049***	52.117***	62.583***	41.012***	50.986***	53.442***	40.413***	60.312***
	(1.494)	(2.931)	(1.717)	(9.705)	(1.433)	(2.429)	(1.639)	(2.059)	(2.183)	(1.647)	(2.246)
Adoption*medium	23.957***	29.832***	21.343***	38.145***	22.886***	31.904***	15.164***	23.868***	23.908***	19.296***	27.096***
	(1.274)	(2.594)	(1.432)	(9.710)	(1.163)	(2.085)	(1.368)	(1.818)	(1.768)	(1.386)	(1.928)
Adoption*long	16.386***	23.778***	12.944***	21.538*	15.971***	22.145***	10.601***	17.310***	14.915***	12.970***	18.500***
	(1.392)	(2.758)	(1.587)	(9.432)	(1.314)	(2.260)	(1.584)	(2.009)	(1.959)	(1.508)	(2.094)
Observations	1,513,824	471,168	1,042,656	105,120	1,408,704	789,600	724,224	783,840	729,984	614,976	898,848
R^2	0.201	0.199	0.201	0.237	0.172	0.208	0.179	0.201	0.201	0.139	0.203
Individual FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year-month FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note:Standard errors are clustered at the individual level *p<0.05; **p<0.01; ***p<0.001

1.5.3 Heterogeneous Treatment Effects on the Treated

We now compute heterogeneous treatment effects on the treated (HTT) by segmenting adopters and matched non-adopters in our sample along a number of dimensions including (1) proximity to airline hub, (2) prior LP mileage balance, (3) adoption in peak travel months, and (4) pre-existing versus newly joining LP members. In addition, we also present an analysis of how treatment effects varied with the amount of initial bonus miles offered to customers. We discuss this aspect separately from the above four heterogeneous effects, because the initial bonus offer may also contain supply-side endogeneity (i.e., the firm may send targeted offers). Table 1.9 shows treatment effects on monthly flight spend by segment, broken into three phases of time. As discussed earlier, we focus on the medium- and long-term effects and their implications for managers. We include tables for heterogeneous effects for other dependent variables in Appendix A.3 and discuss the key highlights in the main text.[12]

Proximity to airline hub

[12] All p-values comparing effect sizes of HTT to the ATT and between pair-wise segments can be found in Appendices A.4 and ??.

We examine how treatment effects differ for card adopters who live in a hub zipcode versus a non-hub zipcode. We define a hub zipcode as any zipcode in a metropolitan area where our focal airline has a hub airport. We are unable to reveal the locations of the hub airports due to the airline wishing to remain anonymous. However, at each of the hub airports of our focal airline, its market share is either highest or second-highest among all airlines serving that airport.[13] 31% (69%) are hub (non-hub) adopters.

From columns 2 and 3 in Table 1.9, we observe that hub adopters increase their flight spend more than non-hub adopters in both the medium- and long- term (the differences are statistically significant). The intuition for this follows from customers in hub airports having more attractive opportunities to fly with the focal airline. That is, more flight destinations may be available with fewer flight stops from hubs. However, we do find a positive treatment effect in the long-term for non-hub adopters (that is half the effect size of hub adopters), suggesting that card adoption also generates lift for those living near non-hub airports of the airline.

Similar to the ATT analysis, most of these flight spend effects are driven by corresponding increases in the number of flights purchased by both segments (see Appendix A.3 for more details). However, we find that spend per flight does increase for non-hub customers but not hub customers in the medium-term.

Correspondingly, hub adopters have a higher lift in LP miles earned than non-hub adopters. In the long term, hub adopters have a higher increase in the number of award flights booked than non-hub adopters. We find no differences in partner redemption behavior change between the two segments.

Prior LP mileage balance

[13] https://www.transtats.bts.gov/

We divide customers into high or low mileage balance segments (prior to adoption) using a threshold of 50,000 miles. Further, the median "initial bonus" offered for credit card adoption is 40,000 miles. In other words, a customer who begins in the low balance segment could leap into a much higher mileage balance by adopting the credit card and fulfilling the conditions to earn the initial bonus. We note that 7% of adopters are in the high balance segment. From columns 4 and 5 in Table 1.9, we do not find statistically significant differences in effect sizes between high and low mileage balance segments.

Similarly, effect sizes are not statistically different for miles earned from flights. However, high mileage balance adopters have a greater increase in award flight redemption. The intuition is that the interaction of higher mileage balances and earning more LP miles on top of that allows for more opportunities to redeem flights (as with a larger number of LP miles, more destinations can be accessed). Interestingly, low mileage balance adopters increase partner redemptions in the long-term while high mileage balance adopters do not. This is likely due to the lower threshold of points needed for partner redemptions as compared to award flight redemptions.

Adoption in peak travel month

It is well established in the air travel industry that peak months of travel (by passenger traffic) are November, December, January, June, July, and August.[14] Since those adopting the card during or just prior to peak travel months (November, December, January, June, July, August) may naturally have a higher expectation of future short-term travel than those adopting the card in off-peak months, we further divide adopters into two segments based on their adoption month - peak or off-peak. We obtain 8,165 peak-season adopters (52%) and

[14]https://scottscheapflights.com/glossary/peak-season

7,604 off-peak adopters (48%). However, from columns 8 and 9 of Table 1.9, we do not find statistical differences between the effects for these two segments.

There are no differences in effect sizes for LP miles earned, award flights redeemed, or partner redemptions between these two segments. Overall, these results suggest the airline does not necessarily have to focus on card adoption at particular points of a calendar year, as adopters from both peak and off-peak months behave similarly across a range of behavioral dimensions.

Pre-existing versus newly joining LP members

We define new LP member adopters as the ones that joined the LP concurrently with card adoption (the airline automatically enrolled a card adopter into the LP if they were not an existing member). Prima facie, pre-existing members who adopt a card may be more knowledgable about the perks and benefits they can leverage with the airline. This segmentation yielded 6,406 new-member adopters (41%) and 9,363 existing-member adopters (59%).

In the medium term, existing-member adopters spend $27 per month more compared to new-member adopters who spend $19 more. In the long term, existing- (new-) member adopters spend $19 ($13) per month more. Thus, existing-member adopters substantially spend more after adoption than new-member adopters. The differences are statistically significant. However, the firm still gains when new-to-the-LP card adopters, which suggests a role of the credit card in switching customers who may be flying with other airlines.

Similar to the ATT analysis, gains in monthly flight spend are driven largely by increases in flight purchases for both segments. Existing-member adopters also have a higher lift in LP miles earned. Existing-member adopters have a higher lift in award flight redemptions than new-member adopters, and vice versa for partner redemptions.

Variation in initial bonus miles in credit card offer

Adopters of an airline credit card can receive a substantial number of initial bonus miles if they meet spending requirements in the initial months. We segment adopters based on their initial bonus miles using the median offer of 40,000 miles in our data. This led to 8,225 high-bonus adopters (52%) and 7,544 low-bonus adopters (48%).

In the medium term, high-bonus adopters increase spend by $32 per month compared to low-bonus adopters whose increase is $15. In the long term, high- (low-) bonus adopters spend $22 ($11) per month more. The medium- and long- term differences between high-bonus and low-bonus adopters are both statistically significant. These differences are driven by differences in the effects on number of flights purchased by each segment, with the high-bonus adopters experiencing a much higher jump in flights purchased. Similarly, the jump in LP miles earned by high-bonus adopters is about double that of low-bonus adopters. High-bonus adopters also have about triple the jump in award flights booked in the medium- and long-term as low-bonus adopters, as well as a higher jump in partner redemptions. It makes intuitive sense that adopters who are given a larger initial bonus will then have more miles to spend on redemptions over time.

We note that the comparison between high-bonus and low-bonus adopters requires more care to interpret than the other dimensions considered earlier. This is because the amount of initial bonus is not simply a customer characteristic but one decided by the firm. To the extent that the firm does not target specific customers with initial bonus offers, the supply side endogeneity concerns may be less. However, we do not have data on targeting algorithms that may have resulted in variation in initial bonus miles offered to individual customers, and this set of results should be viewed with that caution in mind.

1.5.4 Robustness Checks

Our objective in this section is to examine the robustness of our findings. In Section 5.4.1, we consider an analysis in which we replace non-adopters as our control group with later card adopters. That is, those who adopt the credit card in 2016 are considered as the treated group, while those who adopt the card in 2017 are considered the control. We therefore have a shorter time series to consider (since it needs to be truncated before the later adopters actually adopt the card) and repeat the matching exercise described in Section 4 for this new approach. In Section 5.4.2, we replicate our analysis with a log-transformation of dependent variables. In Section 5.4.3, we only consider adopters who did not opt-in for emails from the focal airline, to further understand how such less-engaged adopters may respond to credit card adoption. Finally, in Section 5.4.3, we identify segments of adopters whose post-adoption behaviors suggest that they could have anticipations about future travels, and we exclude these customers from our matched sample to see if the rest of the customers have qualitatively similar effect sizes than those estimated in Table 1.6.

Using late adopters as control group for early adopters

In our main analysis, we match card adopters and non-adopters with similar characteristics using rolling entry matching. As a further check of our analysis we used 2017 adopters as a control (e.g., [29]) instead of non adopters. We use the same matching algorithm detailed in Section 4 and specifically match each 2016 adopter with a similar 2017 adopter that adopted in the same calendar month, to account for seasonality from the supply side. Further, we constrain the end point of the data sample at Dec 31, 2016, so that the 2017 adopters are

Table 1.10: Using 2017 Adopter Cohort as Control Group for 2016 Adopters

	Dependent variable:					
	Flight Spend	Num of Paid Flights	Spend per Flight	Miles Earned thru Flights	Num of Award Flights	Num of Partner Redemptions
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption*short	40.375***	0.193***	0.872	241.201***	0.010***	0.003***
	(2.132)	(0.008)	(6.355)	(14.266)	(0.002)	(0.001)
Adoption*medium	23.568***	0.096***	7.063	152.946***	0.026***	0.004***
	(2.837)	(0.008)	(7.034)	(18.889)	(0.002)	(0.001)
Observations	214,704	214,704	31,695	214,704	214,704	214,704
R^2	0.217	0.263	0.308	0.238	0.073	0.056
Individual FE	Y	Y	Y	Y	Y	Y
Year-month FE	Y	Y	Y	Y	Y	Y

Note:Standard errors are clustered at the individual level *p<0.05; **p<0.01; ***p<0.001

always non-adopters for the purposes of this analysis. Because of the reduced timeframe, we are able to estimate treatment effects for up to 12 months after card adoption, and do not include the long-term time phase (which involves effects over 12 months after adoption).

In Table 1.10, we report short-term and medium-term effects for the set of dependent variables in our study. As previously noted, the short-term effects likely include effects from endogenous timing of adoption, and we do not look to draw conclusions from these effect sizes. However, it is reassuring that the medium-term effect sizes for all of the dependent variables are very similar to the ones reported in Tables 1.6 and 1.7. These findings suggest our key results to be robust to a substantial change in the composition of the control group.

Log transformation of dependent variables

In Table 1.11 below, we report DID estimation results with dependent variables being log-transformed to account for potentially skewed distributions of our dependent variables. The main findings remain unchanged.

Adopters who did not opt-in for emails

In this robustness check, we analyze effect sizes for card adopters who did not opt in for emails from the firm. These adopters seemingly are less engaged in communications (including

Table 1.11: Log Specifications

	Dependent variable:					
	Log (Flight Spend + 1)	Log (# Paid Flights + 1)	Log (Spend per Flight + 1)	Log (# Award Flights + 1)	Log (Miles Earned thru Flights + 1)	Log (# Partner Redemptions + 1)
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption*short	0.664***	0.122***	0.006	0.007***	0.810***	0.002***
	(0.012)	(0.002)	(0.009)	(0.001)	(0.015)	(0.0003)
Adoption*medium	0.244***	0.046***	0.020*	0.023***	0.340***	0.004***
	(0.008)	(0.002)	(0.009)	(0.001)	(0.011)	(0.0002)
Adoption*long	0.153***	0.029***	−0.008	0.023***	0.226***	0.004***
	(0.009)	(0.002)	(0.011)	(0.001)	(0.012)	(0.0002)
Observations	1,513,824	1,513,824	163,182	1,513,824	1,513,824	1,513,824
R^2	0.206	0.227	0.389	0.062	0.216	0.044
Individual FE	Y	Y	Y	Y	Y	Y
Year-month FE	Y	Y	Y	Y	Y	Y
Note:Standard errors are clustered at the individual level						*p<0.05; **p<0.01; ***p<0.001

Table 1.12: Credit Card Adopters Who Did Not Opt-in for Emails

	Dependent variable:					
	Flight Spend	Num of Paid Flights	Spend per Flight	Num of Award Flights	Miles Earned thru Flights	Num of Partner Redemptions
	(1)	(2)	(3)	(4)	(5)	(6)
Adoption*short	32.110***	0.157***	−9.698	0.008***	176.674***	0.004***
	(1.621)	(0.006)	(6.929)	(0.001)	(10.812)	(0.001)
Adoption*medium	14.879***	0.064***	0.607	0.030***	90.486***	0.005***
	(1.475)	(0.005)	(8.596)	(0.001)	(10.857)	(0.0004)
Adoption*long	8.429***	0.038***	−11.276	0.029***	50.317***	0.006***
	(1.607)	(0.005)	(9.015)	(0.001)	(12.837)	(0.0004)
Observations	723,552	723,552	53,444	723,552	723,552	723,552
R^2	0.189	0.249	0.335	0.058	0.233	0.037
Individual FE	Y	Y	Y	Y	Y	Y
Year-month FE	Y	Y	Y	Y	Y	Y
Note:Standard errors are clustered at the individual level						*p<0.05; **p<0.01; ***p<0.001

potential emails relating to credit card offers). The question is whether such adopters do not change much due to the credit card. In Table 1.12, we present the results across the set of dependent variables we study. The effects that were statistically significant in the main analysis continue to hold for this subset of adopters. However, the effect sizes are smaller for this segment.

Alternative selection mechanisms that could drive long-term effects

Although we find that our effect sizes plateaued after 13 months post-adoption, and that forward-looking to the extent of more than two years is extremely difficult, one could still argue that certain customers could anticipate their future travels would go up for many years ahead. For example, say someone recently changed their job that requires more business

Table 1.13: Excluding Adopters Who Could Have Anticipations About Future Travels

	Dependent variable:
	Flight Spend
Adoption*short	45.355 ***
	(1.355)
Adoption*medium	18.942 ***
	(1.181)
Adoption*long	12.278***
	(1.277)
Observations	1,407,552
R^2	0.160
Individual FE	Y
Year-month FE	Y
Note:Standard errors are clustered at the individual level	*p<0.05; **p<0.01; ***p<0.001

travels, or someone recently had some personal life events that require more leisure travels, especially on the same route over time (e.g., going long distance with partners). In these cases, they could foresee their travels would go up for many years ahead. To alleviate this concern, and to make sure that our results are not driven by these alternative selection mechanisms, we first identified segments of adopters who 1. increased the proportion of business travels (round-trips without Saturday night stayovers according to [17]) post-adoption 2. increased the proportion of leisure travels (round-trips with Saturday night stayovers), and with more than 50% of leisure trips with the same origin and destination post-adoption. These two segments account for a total of 1,107 customers, and we report estimation using the rest of the customers in Table 1.13. We still find qualitatively similar results excluding customers who could have anticipations about future travels, suggesting that alternative selection mechanisms are unlikely to have driven our long-term results.

1.6 Discussion

In this study, we estimate the treatment effects of adopting a co-branded credit card on spending and loyalty behaviors with the focal firm. Our data set was obtained from a North American airline firm and contained detailed records of both card adopters and non-adopters. Because customers self-select into adopting a credit card, the adoption decision is not randomized between the treated and control groups. To mitigate selection effects, we use a multi-pronged approach to identifying treatment effects by (1) using a rolling-based matching procedure to obtain covariate balancing for adopters and non-adopters, (2) using difference-in-differences estimation on the matched sample with a two-way fixed effects specification, and (3) dividing effects temporally into short- (1-3 months of adoption), medium- (4 to 12 months after adoption), and long- term (13 months or more after adoption) phases as we argue that short-term effects are most likely to be affected by endogeneous adoption timing, and that long-term effects are least likely to be effected.

While we find large short-term effects on monthly flight spend from card adoption, the challenge in disentangling endogenous adoption timing from actual card effects means that we do not look to draw strong inferences on short-term treatment effects. In the medium- and long- term, flight spend is lifted by 64% and 42% respectively with card adoption. These are statistically significant and economically meaningful effects from card adoption. The firm therefore obtained higher revenues from card adopters even more than a year after adoption. These effects, which have not previously been documented in the literature, suggest there is more to co-branded credit cards than simply earning revenue for selling loyalty program points to financial institutions. Taking the magnitude of these effects into account may help the firm assess the return on investment of such cobranding efforts more comprehensively. Returning to the question of short-term effects, we estimate a gain of $52 in monthly flight spend for adopters, inclusive of endogenous timing effects. It is plausible, given the significant effect sizes for the medium- and long- term phases, that at least some of the short-term gain is also driven by card adoption effects.

We find that increased monthly flight spend is largely driven by an increase in the number of flights purchased by adopters, and not due to adopters paying higher prices for flights. That is, if we assume customers' category-level travel demand is constant, the increase in paid flight bookings would then reflect a higher share of wallet for the airline in our study. This may be driven by economic factors such as the bonus multiplier miles by booking with the focal airline, or the travel perks associated with the card, or may be due to psychological benefits from card adoption

In terms of loyalty program interactions, adopters earn significantly more miles in the loyalty program, in line with the increased flight spend. We also find that adopters increase award flight redemption with the focal airline to a greater extent than partner redemption, suggesting the co-branded credit card focused adopters more on rewards with the firm itself. While loyalty program behaviors may not directly contribute to revenue, higher engagement with the program may help with maintaining customers' relationship with the brand, which may have psychological benefits.

We also find that some types of card adopters are more valuable to the firm in terms of increased spend. Specifically, those adopters who live near hub airports of the focal airline, or who are already existing members of the loyalty program have a higher spend increase in the medium- and long- term. This suggests a co-branded credit card effectively increases the value of customers who already have connections to the firm, either in terms of location or existing participation in the loyalty program. Our finding implies that a firm need not focus their marketing efforts for these cards only on switching customers from competitors.

However, new-to-the-program adopters also experience a lift in spend, and our analysis helps quantify the relative gains from each of these segments.

Finally, we find that adopters with a higher bonus offer do better in terms of increased flight spend than those with a lower bonus offer. If taken at face value, this suggests the use of higher bonus offers to attract card adopters. This, however, should be interpreted with caution as it may include effects of targeting (which are not available in our data set). That is, if higher bonus offers were sent to potential adopters who spend more, it may influence the effects that we find. Experimentation by firms to better understand the optimal level of bonus offers can be valuable in future research.

We believe the above set of findings to be novel to the marketing and loyalty program literature in terms of the effects of co-branded credit cards on customer behavior. We acknowledge that our data is from a single firm, and that any empirical generalizations are only feasible with additional studies across firms and industries. Another fruitful area for future research is to conduct a randomized field experiment in making co-branded credit card offers. This could be done with a firm's existing loyalty program members and provide a complementary approach to measuring effects from an intent-to-treat perspective.

We hope future work can explore the effects of co-branded credit cards across other contexts to expand the body of knowledge in this emerging area of research.

Chapter 2

An Experimental Investigation of Price vs. Non-Price Messaging in Subscription Programs

2.1 Introduction

Subscriptions of digital/content and physical goods are becoming increasingly popular. Much of the increase come from digital content or SaaS,[1] where the subscriptions are about the availability of a service, and the subscriptions are offered by manufacturers/content providers (e.g., Amazon Prime, Spotify, Netflix, etc.) However, subscriptions for physical goods by retailers are very common, too. Examples of non-digital subscription services include product delivery at periodic intervals such as the program offered by dollar shave club, meal prep

[1] https://www.oberlo.com/statistics/most-popular-subscription-services

services like Hello Fresh, or several pet food manufacturers. In this paper, we look at pet products, with pet food as the most-common product.

There are several differences between digital subscriptions, or even other content such as subscriptions for magazines or newspapers, and subscriptions for physical goods. First, it is often the retailers, and not the manufacturers, who are pressing to sell a product by subscription. One consequence is that subscriptions for pet food or air conditioner filters, or other common physical goods compete with retailers selling these items not in a subscription form: If you want to watch Netflix or the New York Times online, you need to subscribe to get this content, but if you want to buy pet food you can buy the product without a subscription from many retailers. You can technically buy a physical New York Times or magazine without a subscription, but note that in these information-based products, it is the manufacturer who is selling the subscriptions – and generally setting the news stand price – so the relative tradeoff of a subscription vs. the piecemeal price is set by the same company. In contrast, we are going to study the case where retailers are offering subscriptions for pet products that can be bought without subscriptions. Second, it is much more important to meter the quality of physical products in a subscription. If you send product too often there are storage issues for the customer, but if you send the product too infrequently then the customer experiences shortages unless they make a trip outside of the subscription to buy more product. In contrast, for digital streaming or SaaS product, there is not even a decision of how frequently to offer products – the spigot is always on. Finally, mistakes in subscription timing for physical goods are very salient. A subscription to Netflix may be a waste of money for people who do not watch it in a particular month, but the misallocation of the subscription is not as obvious and front of mind as it would be for a customer who has 3 bags of dogfood piling up.

From the firm's perspective, the goal of offering subscription is to not only sign-up customers but also to retain them in the longer term. Retention and the recurring revenue it offers is often discussed as one of the key benefits of subscription for firms. Thus, how to attract and retain customers for subscription is of great managerial importance. First of all, subscriptions are often advertised with steep price discounts for customer acquisition. Firms compete heavily on price discounts and might consider it a competitive necessity to offer price discounts. For example, one of the main players in the subscription business, Amazon, gets people to subscribe by having them check a box along with a message to save some amount of money if they subscribe. However, the challenge associated with advertised price discounts is also substantial, as the featured price discounts highlight price savings and this might attract customers signing up for just the price discount and then churning soon after. A possible mechanism to get consumers to focus on motivations other than price is to include non-price-oriented messaging is advertising. This is a novel idea as it focuses on acquisition mechanisms to reduce churn whereas the literature on churn focuses on post-acquisition interventions to reduce churn (e.g., [30]).

In this paper, we examine the relative performance of price versus non-price messaging in attracting and retaining customers for a product subscription service in this paper. We present results from a field experiment where we randomized price and non-price messaging in email advertising of subscription programs. Both the price and non-price advertising offered the same price discounts but in the non-price messaging condition we included other non-price motivations in addition to price (savings). We examine the relative performance of these messages in signing up customers and the subsequent order rates for a subscription service.

A priori, it is not obvious which message will perform better for the firm. Price oriented messages might have a higher sign-up rate due to the salience of the price message and a lower reorder rate whereas the non-price messages might have a lower sign-up rate versus a higher reorder rate. In fact, the firm we worked with believed that price messages were always the ones that worked best and gave a substantial discount on their first orders. We conjecture that our firm is not alone here and that several firms believe that price discounts are most effective message. As we mentioned earlier, even Amazon encourages customers to subscribe by displaying a box along with a message to save money if they subscribe.

Our results show the non-price messages perform slightly better than the price messages in terms of sign-up rates. This is an interesting finding as there is an a priori concern that the non-price messages might crowd out the pricing information and hence mitigate the incentive to sign up because the price discount was made less salient in this case. In addition, we find that the non-price messages outperform price message and the differences between these messages expand over the subsequent orders. This pattern also holds for total number of orders, total revenue and profit margin. We find that the inferior performance of the price message is primarily due to price message attracting lower quality customers. Specifically, we find that price tends to attract a higher proportion of early churners that canceled after one or two orders. Apart from that, price also tends to attract sophisticated gamers with higher levels of gaming behaviors. Specifically, they use credit card tricks (e.g., virtual credit cards) to save time to cancel. Overall, our results show that our firm would be better off with the messages that highlight non-price motivations in addition to just price ones.

To make sure there's a greater match between messages and customers, we use previous purchase history to find the optimal message for each customer, such that customers are less likely to churn later. Our findings suggest that price message should be sent to customers who are less familiar (or have less experience) with the online channel, who are new to subscriptions, who have more regular purchase history (more regular purchase intervals), who are less deal-prone (bought fewer items on sale), and who have purchased fewer private labels. The rest should receive non-price messaging. Moreover, customers with no prior engagement with the retailer should be sent risk messaging. This is likely because they don't have familiarity or they do not trust the retailer, so reassuring them that they will not get items they don't want is important for drawing them in. Finally, those who are most deal-prone and most familiar with online shopping should not receive an email at all. Our research also has implications on how firms that offer subscriptions should advertise to attract customers. It provides insights into the role of messaging in attracting and in retaining customers. More generally the role of messaging and creatives in advertising is an area of growing interest and our work contributes to this. Our work also contributes to the role of advertising messages as an alternative to price competition. Finally, our work suggests that advertising messages can induce self-selection by customers in a manner that can provide a better match and reduce churn.

2.2 Literature Review

2.2.1 Subscription

The first stream of literature that we relate to is subscriptions. Subscription services have grown in popularity and have drawn academic attention in recent years. [22] find that Spotify subscribers increase their music consumption volume, diversity and new music discovery. [31] find that subscription programs lead to significant increase in customer purchase, and that one third of the increase is attributable to economic benefits offered by subscription programs. [32] used a randomized field experiment and find that customers' email engagement can increase service consumption and reduce churn probability. While these papers focus on subscriptions where customers pay a fixed amount of money monthly in order to access program benefits, our paper focuses on subscriptions on a replenishment basis where products are delivered at regular intervals. We add to the growing empirical literature on subscriptions, and we study how email contents affect customer ordering and staying with subscriptions.

2.2.2 The Effect of Email Contents on Customer Response

The second stream of literature that we contribute to is the effect of email contents on customer response. [33] varied ad content along many dimensions (e.g., rate, number of loan options, male vs. female model, etc.) and find that advertising content significantly affects loan take-up rate. [34] investigated how different messages affect charity giving. They randomized advertising content to manipulate sympathy and find that sympathy biases significantly affect donation. [35] study how personalization affects ad response. They added names to email subject lines and find large economic effects on opening and purchase rate. Our paper examines how different email contents affect subscription order rates and retention and our findings suggest that firms are better off using non-price motivations.

2.3 Experiment Setup and Data Descriptions

2.3.1 Experiment Setup

We worked with a major retailer that sells pet products. Understanding what customers are looking for in a pet product subscription is key to attracting and retaining them. Thus, we talked to the retailer and found out five main motivations of customers signing up for subscriptions. First, as we discussed earlier, retailers almost always compete on price such that customers will benefit from price savings if they subscribe (e.g., 35% off first order with max $20 savings, 5% on subsequent orders plus free shipping). Second, consumers also value convenience because products are shipped to them automatically. So there's no lugging around products. Third, having a stable and reliable supply can be especially important for pet owners, as having a subscription can serve as a reminder so people do not forget to buy. Fourth, during the time of our study (1/25/2021 – 2/22/2021), customers also wanted to avoid physical stores due to the risk of contracting COVID. Hence, subscriptions for pet products can provide additional safety. Finally, consumers have fear holding them back from subscriptions: they worry that if they sign up for a subscription they will forget about it and then get a delivery of product they end up not wanting. We tell consumers that they will be informed before a later shipment is made so they have low risk of getting a shipment they do not want.

We ran a four-week experiment from 1/25/2021 – 2/22/2021, during which we randomized price and non-price messaging described above in email advertising of subscriptions. Both the price and non-price advertising offered the same price discounts (35% off first order with max $20 savings) but in the non-price messaging condition a non-price message was prominently featured in addition to the price discount. Specifically, we crafted different messages to match the 5 motivations raised previously: price (35% off first order and free shipping), convenience (products are shipped to you automatically), reminder (having a subscription can serve as a reminder so you do not forget to buy), safety (having a subscription means that you don't have to go to a store, so you are less at risk from COVID), and risk (we tell consumers that they will be informed before a later shipment is made). The experiment also had a control condition where the respondents did not get a subscription solicitation email.[2] However, customers in the control condition could still search and go to the website, and get the first order discount organically at checkout without an email promotion message.

[2]Our experiment did not alter whatever other emails customers received from the retailer (e.g., for non-subscriptions or other products).

What we included in the email subject lines for the five messages match the five motivations discussed earlier, and are as follows: “Save with subscriptions” (price); “Let us do the heavy lifting with subscriptions” (convenience); “Never forget with subscriptions” (reminder); “Safe, no contact shopping with subscriptions” (safety); “No surprises with subscriptions” (risk). The email imagery and the exact phrasing (see Figure 2.1) have been changed to protect the confidentiality of the retailer.

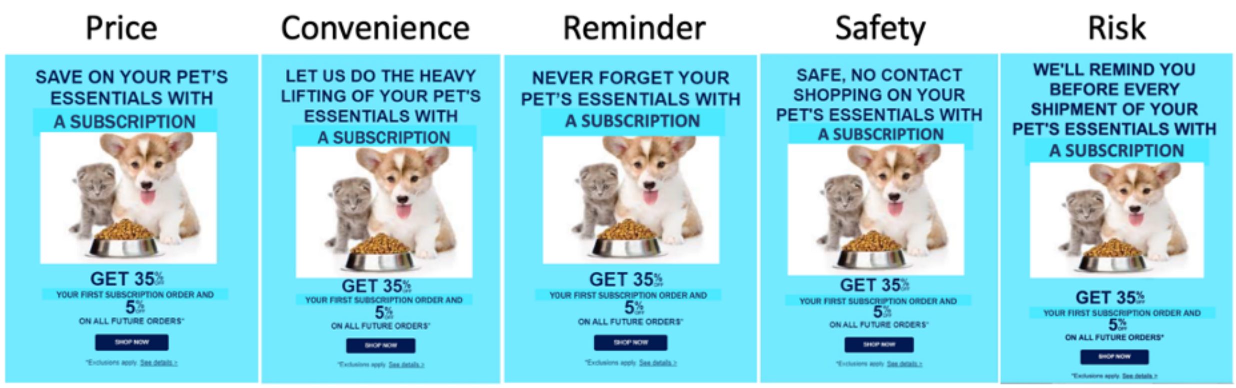

Figure 2.1: Email Illustrations by Condition

The retailer sent out emails soliciting subscriptions once per week to customers who had opened any of this retailer’s email in the past 90 days. Each person got the same message for all 4 weeks. Once people bought a subscription they are removed from the list. In total, 7,928,990 customers were included in our experiment.

2.3.2 Data Descriptions

We have four datasets. The first two datasets contain all customers’ (including those in the control condition) online and in-store transactions from 1/1/2019 to 12/31/2021 (our experiment started on 1/25/2021 and ended on 2/22/2021), including purchase date, SKU, price discount, quantity, sales revenue, Customer ID, and subscription ID (for subscription orders). Note that subscriptions can only be purchased through online channel, and we can know how many times a subscription was ordered from the online order data. The third dataset has information on all subscriptions placed during the same period of online transactions, including Customer ID, subscription ID, SKU, subscription placed date, delivery interval, the trajectory of each subscription's status over time. The status of a subscription is in one of four mutually exclusive states at any point in time: active, canceled, paused, or inactive. Finally, we also observe subscription cancel date if the subscription is cancelled.[3] The last dataset has SKU-level information for all SKUs offered by this retailer, including brand, product department, shipping weight, animal type, etc. In the next two sections we provide evidence for valid randomization and experiment outcomes by condition.

2.4 Randomization Check

To confirm that we have valid randomization between conditions, we calculate six customer purchase behaviors between online (Ecom) and in-store (BM) channels in the past 90 days when email targeting criteria took place, and report the results in Figure 2.2 below.

At first glance, different conditions do have similar prior purchase behaviors between online (Ecom) and in-store (Brick and Mortar) channels. We also conduct formal pairwise t-test for every purchase metric. We find that 6/90 (15*6) pairs are statistically different at 10% level, and 0/90 pairs are statistically different at 5% level. Thus, we confirm that groups are similar before experiment, and that we have valid randomization.

[3]if a subscription is canceled, it cannot return to another status.

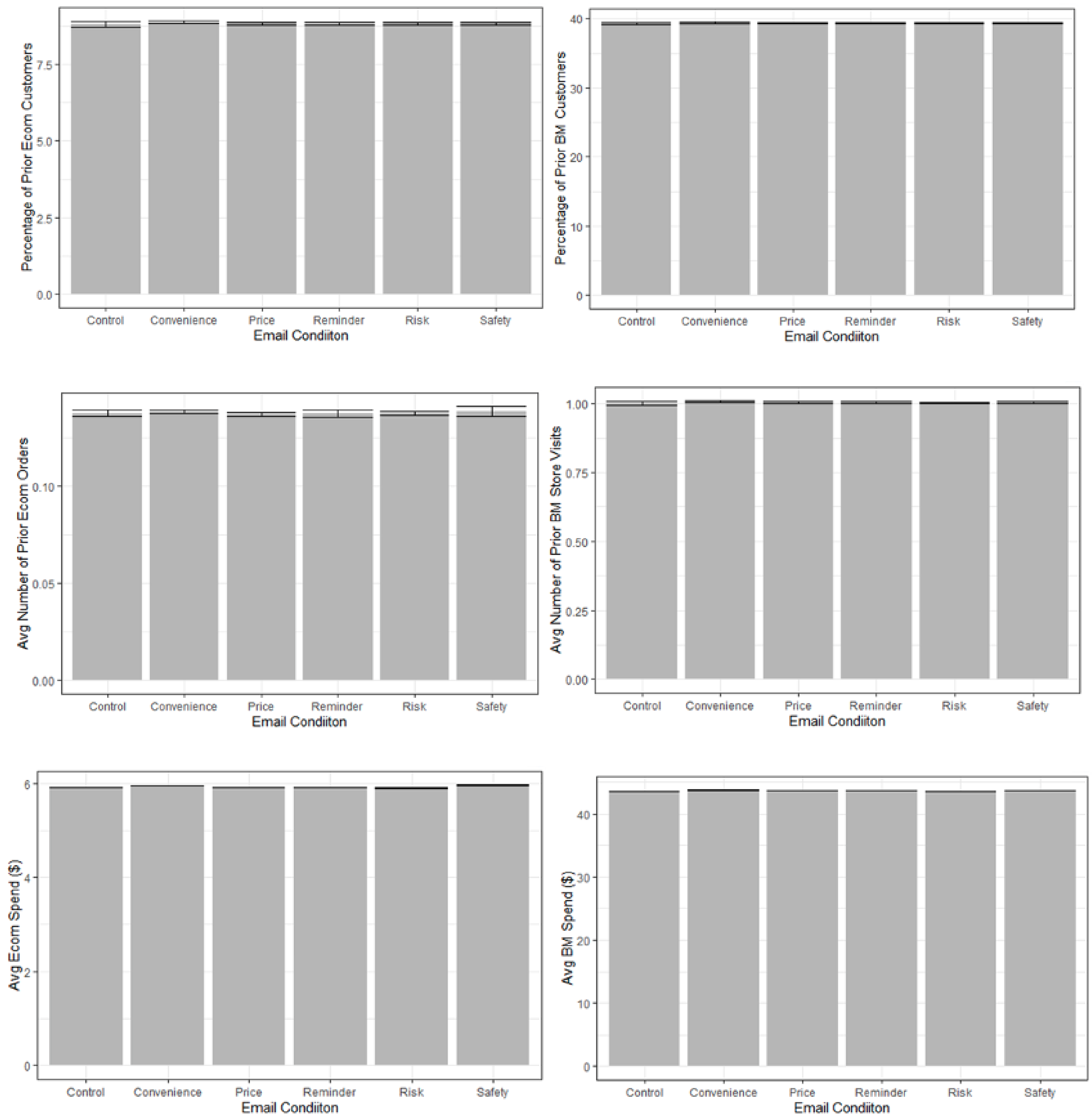

Figure 2.2: Randomization Check

2.5 Experiment Outcomes

2.5.1 Email Open and Click-through Rate

To examine the relative performance of price and non-price messages, we start with email open and click-through rate.

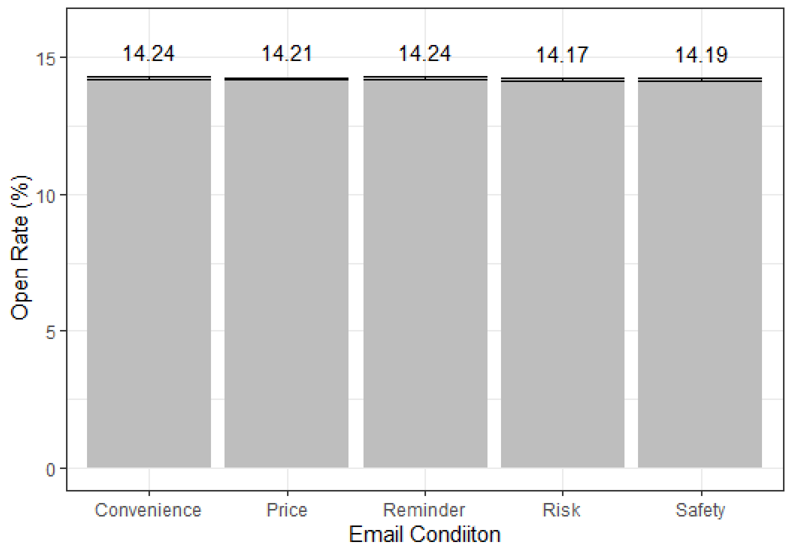

Figure 2.3: Email Open Rate by Email Condition

From Figures 2.3 and 2.4, we can see that price gets some initial attention with clicks (price is not statistically different from other conditions at 10% level for opens), and is statistically different from the convenience and the safety messages at 10% level. This might explain why the retailer believed price worked well. On the other hand, non-price messages do not do that much worse. This was one of the firm's concerns that non-price would not get attention. Next, we will show that once we get to orders, price does not perform well compared to non-price messages

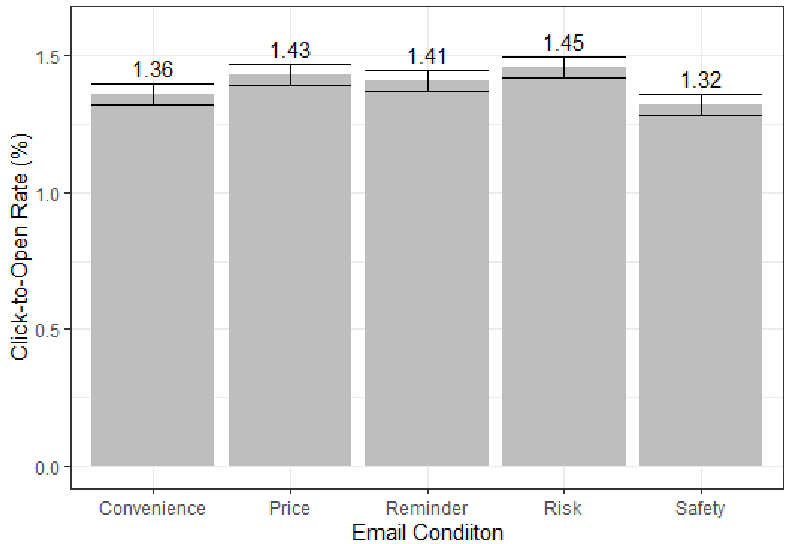

Figure 2.4: Email Click-through Rate by Email Condition

2.5.2 First, Second, and Third Order Rates by Email Condition

A natural starting point is to look at probability of sign-ups and reorders for at least one subscription. Our retailer primarily offers four product categories: companion animals (e.g., products for animals other than cats and dogs, like fish or hamsters), consumables (e.g., pet food), dog and cat supplies (e.g., dog and cat furniture), and RX/prescription. In total, 34,202 subscriptions were placed by 21,287 customers during the 4-week period. Specifically, 75% of subscriptions placed across conditions belong to consumables, and among which, around 70% of consumables are dog and cat food. Companion animals account for 11%, and dog and cat supplies account for another 11%. Finally, RX/prescription accounts for the rest 3% of total subscriptions.

We first provide first, second and third order rates by email condition in Table 2.1. The first order rate is defined as the percentage of customers in experiment who ordered at least one subscription during the 4-week experimental period. Starting from the second order, customers could choose to reorder (or not) any of their subscriptions. So we define the second and third order rates as the percentage of customers in experiment who reordered a second and a third time for at least one subscription placed during the experimental period.[4] As we noted earlier, even though customers from the control condition did not receive a solicitation email, they could still go to the website and get the first-order discount at checkout organically.

Table 2.1: First, Second and Third Order Rates by Email Condition

	Control	Convenience	Price	Reminder	Risk	Safety
Number of Customers	418,925	1,502,035	1,501,944	1,502,090	1,502,084	1,501,912
First Order Rate	0.2561%	0.2717%	0.2626%	0.2750%	0.2690%	0.2675%
	(0.0078%)	(0.0043%)	(0.0043%)	(0.0042%)	(0.0042%)	(0.0042%)
	[0.48]	[0.13]		[0.04]	[0.29]	[0.41]
Second Order Rate	0.1618%	0.1703%	0.1613%	0.1686%	0.1696%	0.1649%
	(0.0062%)	(0.0034%)	(0.0033%)	(0.0034%)	(0.0034%)	(0.0033%)
	[0.95]	[0.06]		[0.12]	[0.08]	[0.44]
Third Order Rate	0.1067%	0.1155%	0.1058%	0.1134%	0.1146%	0.1124%
	(0.0050%)	(0.0028%)	(0.0027%)	(0.0028%)	(0.0028%)	(0.0027%)
	[0.89]	[0.01]		[0.05]	[0.02]	[0.09]

Note: Standard Errors in Parentheses and P-values vs. Price in Brackets

Since our goal is to compare the relative performance of non-price messages and price message, we focus on the five email conditions and will detail how non-price messages outperform price in terms of order rates. For the first order in Table 2.1, we find that the differences across email conditions are only statistically significant between price and reminder (reminder slightly outperformed price at 10% significance level), and are statistically indistinguishable across all other conditions. Reminder gives a lift of 4.7% relative to price for first order rate. Since the retailer makes very little profit on first order with 35% off (max $20 savings), it is also important to understand which message is most effective at retaining customers for longer term. We further calculate percentage of customers that reordered a second and a

[4]By definition, customers who reordered a third time are a subset of those who reordered a second time.

third time. We find that the differences across conditions expand and become more significant. Specifically, we find that all four email messages outperform price and the differences are all statistically significance at 10% level when we reach third order. Non-price messages give a lift of 6.5% - 10% for third order rate relative to price. The superior performance of non-price messages holds across all product categories.

2.5.3 Total Number of Orders, Revenue and Profit Margin

In this section, we show that non-price messages are better than price across different metrics. We discuss our robust findings along multiple dimensions including total number of orders, revenue and profit margin.

Given that customers could order multiple subscriptions during the experiment, a more relevant statistic than order rates would be the total number of subscription orders, which we report in Table 2.2 and Figure 2.5.[5]

Table 2.2: Avg Number of Orders per 100 People Emailed

	Control	Convenience	Price	Reminder	Risk	Safety
Number of Customers	418,925	1,502,035	1,501,944	1,502,090	1,502,084	1,501,912
Avg num of orders at first order	0.4168	0.4278	0.4231	0.4411	0.4265	0.4424
per 100 people emailed	(0.0154)	(0.0081)	(0.0082)	(0.0084)	(0.0083)	(0.0103)
	[0.73]	[0.25]		[0.05]	[0.31]	[0.01]
Avg num of orders up to third order	0.7913	0.8108	0.7910	0.8253	0.8094	0.8294
per 100 people emailed	(0.0304)	(0.0158)	(0.0160)	(0.0162)	(0.0161)	(0.0184)
	[0.49]	[0.08]		[0.01]	[0.1]	[0.01]
Avg num of orders up to fifth order	0.9801	1.0095	0.9774	1.0181	1.0037	1.0291
per 100 people emailed	(0.0394)	(0.0207)	(0.0211)	(0.0211)	(0.0210)	(0.0239)
	[0.47]	[0.06]		[0.02]	[0.98]	[0.01]

Note: Standard Errors in Parentheses and P-values vs. Price in Brackets

[5] We bootstrapped standard errors 10k times for each email condition, and we then calculate p-values based on a simulation study where we count % of times non-price messages are larger than price message.

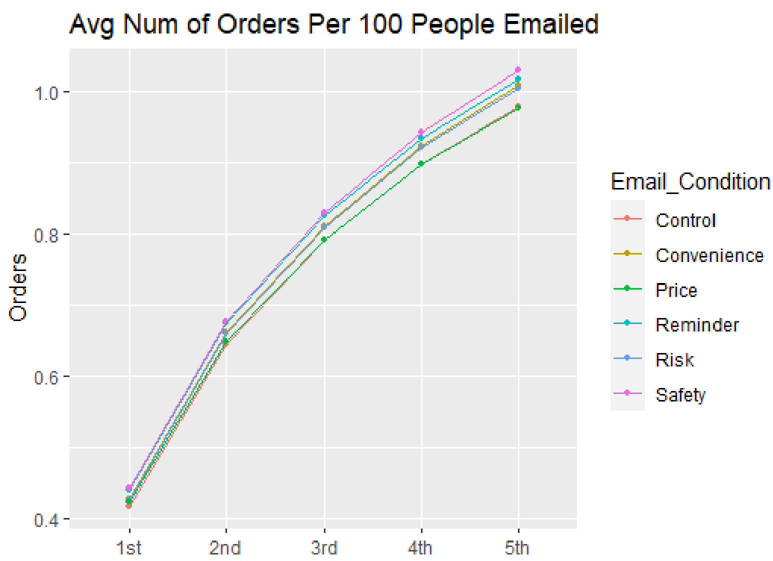

Figure 2.5: Average Number of Orders per 100 people Emailed

We find a similar pattern to what we see in order rates, which is that the differences between price and non-price conditions (in terms of number of orders) are small at first order, but gradually expand and really show in subsequent orders. Specifically, safety and reminder perform the best at first order ($p < 0.012$). And this difference grows even more as we reach third and even push to fifth order. Up to fifth order, all email messages are statistically different from price at 10% level, and Figure 2.6 shows that all non-price conditions are above price at fifth order. The best messages (safety and reminder) in general are about 4-5% better than price.

From the firm's perspective, it is managerially meaningful to also look at retention in terms of revenue. Our finding that non-price messages outperform price message holds for revenue, which we report in Table 2.3 and Figure 2.6. We see that price is consistently lower than other non-price messages. Reminder and convenience are both statistically different from price at 10% level from first order to fifth. Figure 2.6 demonstrates that the differences expand as we push further in subsequent orders. The best messages (reminder and convenience) are about 4-6% higher than price.

Table 2.3: Avg Revenue per 100 People Emailed

	Control	Convenience	Price	Reminder	Risk	Safety
Number of Customers	418,925	1,502,035	1,501,944	1,502,090	1,502,084	1,501,912
Avg revenue at first order	10.85	11.46	11.41	12.11	11.21	11.54
per 100 people emailed	(0.42)	(0.22)	(0.23)	(0.27)	(0.22)	(0.28)
	[0.92]	[0.06]		[0.05]	[0.66]	[0.24]
Avg revenue up to third order	24.12	25.83	25.14	26.69	25.17	25.64
per 100 people emailed	(0.94)	(0.53)	(0.53)	(0.57)	(0.51)	(0.56)
	[0.93]	[0.08]		[0.001]	[0.48]	[0.17]
Avg revenue up to fifth order	31.03	33.44	32.21	34.19	32.43	33.06
per 100 people emailed	(1.29)	(0.73)	(0.73)	(0.79)	(0.70)	(0.75)
	[0.87]	[0.04]		[0.006]	[0.38]	[0.13]

Note: Standard Errors in Parentheses and P-values vs. Price in Brackets

Readers might also be interested in profit margin across conditions. We want to note that our margin data is very coarse because ideally the margin should be at the UPC level. However, our margin data is at the brand-department level (e.g., blue buffalo – dog food). This would bring additional noise to our profit calculation given the small effect sizes in our experiment. We present the profit margin by condition in Table 2.4 below.

Table 2.4: Avg Margin per 100 People Emailed

	Control	Convenience	Price	Reminder	Risk	Safety
Number of Customers	418,925	1,502,035	1,501,944	1,502,090	1,502,084	1,501,912
Avg margin at first order	0.53	0.52	0.53	0.55	0.44	0.54
per 100 people emailed	(0.07)	(0.04)	(0.04)	(0.05)	(0.04)	(0.07)
	[0.54]	[0.62]		[0.43]	[0.96]	[0.50]
Avg margin up to third order	2.98	3.18	3.05	3.15	2.97	3.13
per 100 people emailed	(0.19)	(0.11)	(0.10)	(0.11)	(0.10)	(0.12)
	[0.64]	[0.17]		[0.21]	[0.73]	[0.29]
Avg margin up to fifth order	4.13	4.52	4.27	4.42	4.23	4.46
per 100 people emailed	(0.25)	(0.16)	(0.14)	(0.16)	(0.14)	(0.16)
	[0.72]	[0.09]		[0.21]	[0.58]	[0.17]

Note: Standard Errors in Parentheses and P-values vs. Price in Brackets

In general, we still see that non-price messages perform better than price for profit margin, but our results lack statistical significance due to the nature of our margin data. However,

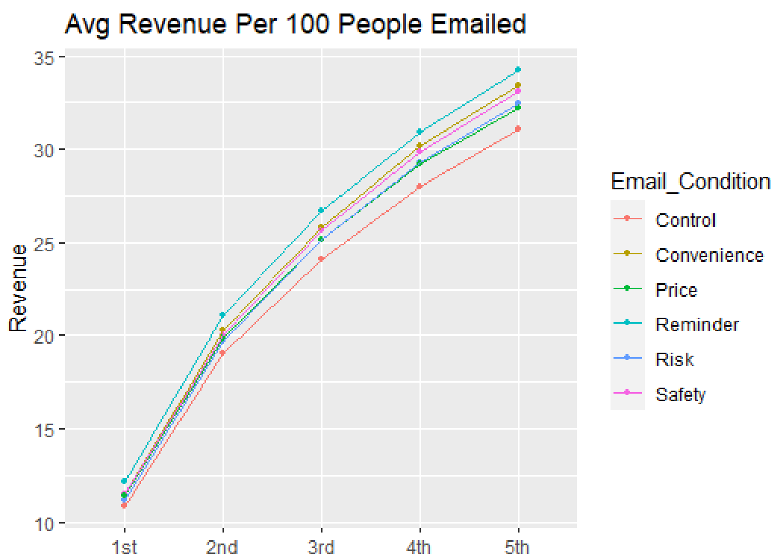

Figure 2.6: Average Revenue per 100 People Emailed

taken together, we find that non-price messages are best for difference metrics. In the next section, we explain why this is the case and show that price and non-price conditions tend to attract different types of customers.

2.6 Accounting for Inferior Performance of Price

2.6.1 Cancellation

From the previous section we see that price was consistently outperformed by non-price messages in terms of sign-up rate, reorder rate, number of orders, revenue, and profit margin. In this section, we seek to explain the selection mechanism of the price condition that contributes to its inferior performance. We will show that price tends to attract a higher proportion of 1. early churners who cancelled within one or two orders 2. and strategic customers who use virtual credit cards. Together they contribute to low order rates and revenue/margin from price.

We first discuss evidence of early cancellations as cancellation is the most common action one could think of when they want to terminate their subscriptions. Thus, we first look at the overall cancellation rate by condition. Since customers that placed multiple subscription orders could choose to cancel any of them, we calculate the percentage of customers who eventually cancelled at least one subscription by 12/31/2021[6] out of all customers who ordered subscription during the 4-week experimental window in Table 2.5. As we explained, one customer could place multiple subscriptions and choose to cancel or reorder any of them, we also examine % of subscriptions that eventually got cancelled by 12/31/2021, and we report robustness in Table B.2 in Appendix B.

We find that the differences in overall cancellation rate across conditions are all statistically insignificant at 10% level, suggesting that the percentage of people who eventually canceled at least one of their subscriptions by 12/31/2021 is similar across conditions. The conclusion is also consistent at the subscription level, i.e., the percentage of subscriptions that got cancelled by 12/31/2021 is also similar across conditions.

Table 2.5: Overall Cancellation Rate

	Control	Convenience	Price	Reminder	Risk	Safety
Number of Signed up Customers	1,073	4,081	3,944	4,131	4,040	4,018
Overall Cancellation Rate	70.64%	72.65%	72.31%	72.45%	72.90%	72.20%
	(1.39%)	(0.70%)	(0.71%)	(0.70%)	(0.70%)	(0.71%)
	[0.30]	[0.75]		[0.91]	[0.58]	[0.93]

Note: Standard Errors in Parentheses and P-values vs. Price in Brackets

If the overall cancellation rate is similar across conditions, why would price consistently have bad performance? It could be that among those who eventually cancelled by 12/31/2022,

[6]Over 90% of the subscriptions have either been ordered three times or more, or been canceled by 12/31/2021.

price has higher percentage of customers who cancelled within one or two orders, a.k.a. early churners, leading to lower second and third order rates. Thus, we also investigate cancellation timing conditional on cancellation. We do so by calculating the percentage of customers who cancelled at least one subscription within one or two orders out of customers who eventually cancelled by 12/31/2021. This analysis has a different focus than what we discussed earlier because we are looking at cancellation more closely here. Specifically, we are interested in cancellation timing conditional on cancellation. We also replicate our analysis at the subscription level, i.e., percentage of subscriptions cancelled within one or two orders, which we report in Table B.3 in Appendix B.

Table 2.6: Conditional Cancellation Rate

	Control	Convenience	Price	Reminder	Risk	Safety
Number of Cancelled Customers	758	2,965	2,852	2,993	2,945	2,901
Conditional Cancellation Rate within One Order	53.56%	51.70%	53.44%	52.52%	50.97%	52.95%
	(1.81%)	(0.92%)	(0.93%)	(0.91%)	(0.92%)	(0.93%)
	[0.98]	[0.19]		[0.50]	[0.06]	[0.73]
Conditional Cancellation Rate within Two Orders	78.10%	76.49%	79.00%	76.34%	76.10%	77.01%
	(1.50%)	(0.78%)	(0.76%)	(0.78%)	(0.79%)	(0.78%)
	[0.63]	[0.02]		[0.02]	[0.01]	[0.07]

Note: Standard Errors in Parentheses and P-values vs. Price in Brackets

In Table 2.6, we find that despite all groups having similar proportions of customers who eventually cancelled by 12/31/2021, price has the highest proportion of early churners (e.g., customers who cancelled after first two orders). Specifically, for conditional cancellation within one order, the differences are statistically significant between price and risk at 10% level (price about 5% higher than risk). For conditional cancellation within two orders, the differences expand and become more significant. All groups are statistically different from price, and price has the highest proportion of early churners (price about 4% higher than non-price messages). The same is true with subscription-level cancellation rate, which we report in Table B.4 in Appendix B. Higher proportions of early churners from price directly contribute to low performance along reorder rate, number of orders, revenue and margin, and can be evidence of price attracting lower quality customers who just wanted a quick deal.

2.6.2 Credit Card Gaming Activities

Apart from early cancellation, we also identified a specific type of gaming behavior in the population. As we noted earlier, the status of a subscription can become inactive, and the inactive status can be triggered by pending RX/prescription approval or credit card failures. What draws our attention is when non-RX/prescription subscriptions became inactive due to credit card failures after just one order and stayed inactive till 12/31/2021. These subscriptions were never cancelled but were not reordered either. A subscription that stayed inactive for more than 10 months after the first order is hard to go unnoticed by a customer (since majority of the products are pet food) and simply cannot be explained by negligence. As a matter of fact, more and more people are using virtual credit cards to order subscriptions these days. The way these virtual credit cards work is that customers can set a total spending limit for a specific merchant, and they do not need to worry about not cancelling subscriptions in time. We believe the subscriptions that became inactive due to credit card failures within just one order were direct evidence of people using these credit card tricks to game the system.

We next present inactive rate due to credit card failures after just one order in Table 2.7. We also make sure we only focus on non-RX/prescription subscriptions that became inactive within just one order, and stayed inactive until 12/31/2021. Because an inactive status can be triggered by pending RX/prescription approval or credit card failure. By just looking at the non-prescription subscriptions, we make sure that the inactive statuses could only be triggered by credit card failures. Specifically, we calculate percentage of non-RX/prescription subscriptions that became inactive after first order and stayed inactive till 12/31/2021. And we present our results in Table 2.7.

Table 2.7: Inactive Rate after First Order

	Control	Convenience	Price	Reminder	Risk	Safety
Number of non-RX Subscriptions	1,693	6,260	6,213	6,455	6,271	6,487
Inactive Rate within One Order	1.95%	2.11%	2.51%	2.35%	2.04%	2.05%
	(0.34%)	(0.18%)	(0.20%)	(0.19%)	(0.18%)	(0.18%)
	[0.21]	[0.15]		[0.61]	[0.09]	[0.09]

Note: Standard Errors in Parentheses and P-values vs. Price in Brackets

From Table 2.7, we find that price has significantly higher inactive rate after first order than risk and safety at 10% level, and price is around 23% higher on credit card gaming activities than non-price messages in general. Note that customers still need to log into their accounts (or call customer service) to cancel their subscriptions. Using a virtual credit card seems like an easier way to not get another shipment after first order without all the hassles. This way, people who use virtual credit cards do not even need to do anything after first order as the virtual credit cards simply would not go through for later orders. Our results show that price attracts a higher proportion of gamers who use these credit card tricks. The benefit of gaming more subscriptions is obviously reselling. Reselling items is something that happens with product subscriptions as opposed to digital services (where sharing of passwords might be the adverse selection problem). Perhaps they are doing this across lots of products (e.g., semiprofessional sellers on eBay) and hence have to use signals of price to buy quickly, and the non-price advertising does not provide that signal.

In addition, we also find that customers who respond to price condition tend to purchase slightly more expensive products. Perhaps the idea is to get the best deal with 35% off and capped at $20 savings. This is also consistent with other evidence that price attracts lower quality customers that do not stay with the firm.

To conclude, in this section, we explain why non-price messages outperform price across different metrics. We find that the inferior performance of the price message is primarily due to price attracting lower quality customers. Price tends to attract a higher proportion of early churners that canceled after one or two orders. Apart from that, price also tends to attract sophisticated gamers who used virtual credit cards to save time to cancel. Overall, our results show that our firm would be better off with non-price motivations in attracting and retaining customers.

2.7 Email Targeting Using Previous Purchase History

Marketing literature has acknowledged the importance of using previous purchase history to better target customers (e.g., [36], [37]). While our experiment randomly sent email messages to customers, we ask two questions in this section: 1. Who should be sent which message if we were to use their previous purchase history for a better match? 2. How much could optimal targeting further lift revenue? The answers to these questions would help us understand how different email messaging appeals to different segments as well as the power of personalization.

We use four sets of explanatory variables to capture prior purchase history from 1/25/2019 – 1/24/2021. Our firm is a multi-channel firm. So the appeal to different customer types across channels is also of interest.

The first set measures prior channel experience. Do subscriptions attract customers who already shop online? An Ecom customer probably sees more messages for subscription every time they shop online. They might also be getting more emails that are triggered by their shopping activities. Thus, there might be a higher exposure and awareness effect. In comparison, a BM customers might have less awareness of subscription services and might have been subjected to fewer marketing messages. They might be less familiar with Ecom and related quality issues and could have concerns about leaving products on the porch. Prior Subscription experience might also be a useful indicator of interest and retention. If the customer had a positive prior experience, then it bodes well for future subscription sign up and retention. In addition, if a customer purchased from the retailer on a regular basis in the past, the demand for a subscription where products get delivered on a regular interval may be higher. Thus, we include whether a customer had purchased subscriptions before, whether a customer had purchased more than or equal to four times from Ecom and BM respectively,[7] and if so, their purchase irregularity, defined as the standard deviation of interpurchase time divided by the mean of interpurchase time.[8]

The second set captures prior shopping preferences/habits. Some customers might primarily shop online and for them online shopping is more convenient and thus product subscriptions could be a greater match. Other customers might primarily shop at BM locations and converting them to subscriptions could be more difficult. We measure prior shopping preference using the number of prior Ecom orders divided by the total number of prior orders.

The third set captures promotion and sale orientation. To remind the readers, our retailer offers 35% off first subscription order and 5% off subsequent orders. Subscriptions that offer a constant price for subsequent orders might be less suited for a buyer who is searching for deals. So prior tendency to purchase on deals might be a marker that we might get lower retention.

[7]We used four purchases as cutoff so that we can more reliably calculate purchase irregularity with at least three interpurchase intervals.

[8]We capped interpurchase time at 180 days/6 months. The results are robust when capping interpurchase time at 90 days/3 months.

Note the Ecom buyers/primarily Ecom buyers might also be deal seekers as search costs are lower online. Moreover, past private label purchases could also indicate sale orientation as the average unit price of private labels is much lower than that of national brands in our data. So we might see lower retention from these customers. Ultimately, we measure deal proneness as the percentage of items bought on sale and the percentage of private labels purchased from both Ecom and BM.

The fourth set measures recency of purchase. A customer who has been dormant for longer might be harder to win back than a customer who had made a more recent purchase. We operationalize recency by measuring days since last Ecom and BM purchase (log transformed) for each customer. Together, these four sets of purchase history variables are apriori relevant both from the perspective of the customer and the retailer.

We now return to the two questions raised earlier in the beginning of this section: 1. Who should be sent which messaging if we were to use their previous purchase history for a better match? 2. How much could optimal targeting further lift revenue? To answer these questions, we use causal forest ([38]), a causal inference learning method to discover heterogeneity and achieve optimal policy evaluation. Specifically, for each customer i, we find the email messaging j that maximizes:

$$CATE_{ij} \coloneqq E[Y_i(j) - Y_i(control)|X_i], \tag{2.1}$$

where $Y_i(j)$ and $Y_i(control)$ represent potential outcome (e.g., \$ revenue through 12/31/2021) under email assignment j and control. X_i represents a high-dimensional vector of prior purchase history discussed previously.

Our results from causal forest suggest that under optimal treatment assignment (everyone receives their best message) 39% of customers should be assigned to the risk messaging, including all customers with no prior engagement (29% in total). We conjecture that customers with no prior engagement with the retailer probably do not trust the retailer. So reassuring them that they will not get items they do not want is important for drawing them in and having them stay longer. In addition to the risk messaging, 12.8% of total customers should be assigned to the convenience messaging, 13.1% to the price messaging, 13.2% to the reminder messaging, 12.1% to the safety messaging, and 9.8% to control.[9]

To understand the value of optimal targeting, we conduct the following counterfactuals. We use the blanket price (i.e., everyone receives the price messaging) as the baseline since price was used by our retailer as the dominant strategy to sell subscriptions. What we want to understand is how much the best blanket strategy (i.e., the reminder messaging) and optimal targeting (everyone receives their best) can further lift revenue. Our results indicate that the blanket reminder can lift the revenue by 5.5% relative to the blanket price. The 5.5% lift in revenue is economically and managerially meaningful as the only thing we change is the email message, which is easily doable by the retailer and could have a large impact on their business. In addition, we also find that optimal targeting can lift revenue by 107% relative to the blanket price. The counterfactuals suggest that the most dominant strategy that firms use to sell subscriptions is very suboptimal, and the effect of email personalization is also huge.

With each customer's optimal assignment, it's also informative to get a sense of which types of customers should be sent price vs. non-price messages. Thus, we compute the group mean of each purchase history variable under optimal assignment.[10] We report the group means in more detail in Table B.4. Our analysis suggests that if the firm wants to maximize revenue,

[9]To understand the "true" impact of email advertising, we exclude 548 credit card gamers we identified in section 6.2. and use the rest of the customers for estimation.

[10]We only do this on customers with non-zero purchase history as customers with zero history should be targeted with the risk messaging by default.

price message should be sent to customers who are less familiar with the online channel, customers who are new to subscriptions(no previous subscription purchases), customers who have more regular purchase history (more regular purchase intervals), and customers who are less deal-prone (bought fewer items on sale/bought fewer private labels). These segments should be targeted with the price messaging likely because those who are less familiar with the online channel, those who are new to subscriptions, and those who are less deal-prone are not as good at finding deals elsewhere online. The price message is less likely to lead to opportunistic behaviors. Similarly, those who have more regular purchase intervals are also less likely to be good deal finders as it is hard to predict when deals will come around (except for holidays).

We also find that customers who are most deal-prone and most familiar with online channel should be not sent an email at all. It's possible that these customers have experience with online shopping and are good at finding deals. Thus, they could always find their way to the website even without a subscription promotion email. In addition, getting deal-prone customers means lower retention for the company. Hence not targeting this segment is better for the business. The rest of the customers should be targeted with non-price messages.

2.8 Conclusions

Taken together, our findings suggest that non-price messages outperform price message along multiple dimensions. The inferior performance of price is due to price attracting lower quality customers who cancel earlier and use virtual credit cards, which prevents better retention. Thus, the retailer is better off speaking to non-price motivations to attract and retain customers. This finding suggests that one of the most dominant ways of selling subscriptions is very suboptimal. Further, firms could also use previous purchase history to better target customers who could be a good match for the subscription services. Our results suggest that price message should be sent to customers who are less familiar with the online channel, customers who are new to subscriptions, customers who have more regular purchase history, and customers who are less deal-prone. The rest should be sent non-price messages. However, customers with no prior engagement with the firm should be targeted with the risk messaging by default. Finally, those who are most deal-prone and most familiar with online channels should not be sent any messages at all. Our research has implications on how firms that offer subscriptions should advertise to attract customers. Specifically, we find that firms are better off speaking to non-price motivations or use previous purchase history to find customers that are a better match. We also provide insights into the role of messaging in attracting and in retaining customers. More generally the role of messaging and creatives in advertising is an area of growing interest and our work contributes to this. Our work also contributes to the role of advertising messages as an alternative to price competition.

References

[1] Jeff Berry. “The 2015 Colloquy loyalty census: Big numbers, big hurdles.” In: *Colloquy Loyalty Talks* (2015).

[2] Pedro Gardete and James Lattin. “Coalition loyalty program not working? Perhaps you’re doing it wrong.” In: *Stanford University Graduate School of Business Research Paper* (2018).

[3] Valeria Stourm, Eric Bradlow, and Peter Fader. “Market positioning using cross-reward effects in a coalition loyalty program.” In: *HEC Paris Research Paper No. MKG-2017-1242* (2017).

[4] Donald Ngwe, Michael Els, Sunil Gupta, and Daniel Schwam. “Cross Merchant Spillovers in Coalition Loyalty Programs.” In: *Working Paper* (2021).

[5] Wayne Taylor and Xiaojing Dong. “Controlling for Retailer Synergies when Evaluating Coalition Loyalty Programs.” In: *Working Paper* (2022).

[6] Arun Gopalakrishnan, Zhenling Jiang, Yulia Nevskaya, and Raphael Thomadsen. “Can non-tiered customer loyalty programs be profitable?” In: *Marketing Science* 40.3 (2021), pp. 508–526.

[7] Yuping Liu. “The long-term impact of loyalty programs on consumer purchase behavior and loyalty.” In: *Journal of marketing* 71.4 (2007), pp. 19–35.

[8] Steven M Shugan. “Brand loyalty programs: are they shams?” In: *Marketing Science* 24.2 (2005), pp. 185–193.

[9] Wesley R Hartmann and V Brian Viard. “Do frequency reward programs create switching costs? A dynamic structural analysis of demand in a reward program.” In: *Quantitative Marketing and Economics* 6.2 (2008), pp. 109–137.

[10] Wayne Taylor and Brett Hollenbeck. “Leveraging loyalty programs using competitor based targeting.” In: *Quantitative Marketing and Economics* (2021), pp. 1–39.

[11] Paul Klemperer. “Competition when consumers have switching costs: An overview with applications to industrial organization, macroeconomics, and international trade.” In: *The review of economic studies* 62.4 (1995), pp. 515–539.

[12] Byung-Do Kim, Mengze Shi, and Kannan Srinivasan. "Reward programs and tacit collusion." In: *Marketing Science* 20.2 (2001), pp. 99–120.

[13] Kimon Proussaloglou and Frank Koppelman. "Air carrier demand." In: *Transportation* 22.4 (1995), pp. 371–388.

[14] Kimon Proussaloglou and Frank S Koppelman. "The choice of air carrier, flight, and fare class." In: *Journal of Air Transport Management* 5.4 (1999), pp. 193–201.

[15] Stephane Hess, Thomas Adler, and John W Polak. "Modelling airport and airline choice behaviour with the use of stated preference survey data." In: *Transportation Research Part E: Logistics and Transportation Review* 43.3 (2007), pp. 221–233.

[16] Mara Lederman. "Do enhancements to loyalty programs affect demand? The impact of international frequent flyer partnerships on domestic airline demand." In: *The RAND Journal of Economics* 38.4 (2007), pp. 1134–1158.

[17] A Yeşim Orhun, Tong Guo, and Andreas Hagemann. "Reaching for gold: Frequent-flyer status incentives and moral hazard." In: *Marketing Science* (2022).

[18] Elizabeth A Stuart. "Matching methods for causal inference: A review and a look forward." In: *Statistical science: a review journal of the Institute of Mathematical Statistics* 25.1 (2010), p. 1.

[19] Paul R Rosenbaum and Donald B Rubin. "The central role of the propensity score in observational studies for causal effects." In: *Biometrika* 70.1 (1983), pp. 41–55.

[20] Shunyao Yan, Klaus M Miller, and Bernd Skiera. "How Does the Adoption of Ad Blockers Affect News Consumption?" In: *Journal of Marketing Research* (2022).

[21] Nathan Blascak and Anna Tranfaglia. "Decomposing Gender Differences in Bankcard Credit Limits." In: (2021).

[22] Hannes Datta, George Knox, and Bart J Bronnenberg. "Changing their tune: How consumers' adoption of online streaming affects music consumption and discovery." In: *Marketing Science* 37.1 (2018), pp. 5–21.

[23] Allison Witman, Christopher Beadles, Yiyan Liu, Ann Larsen, Nilay Kafali, Sabina Gandhi, Peter Amico, and Thomas Hoerger. "Comparison group selection in the presence of rolling entry for health services research: Rolling entry matching." In: *Health services research* 54.2 (2019), pp. 492–501.

[24] David R Bell, Santiago Gallino, and Antonio Moreno. "Customer supercharging in experience-centric channels." In: *Management Science* 66.9 (2020), pp. 4096–4107.

[25] Peter C Austin. "Using the standardized difference to compare the prevalence of a binary variable between two groups in observational research." In: *Communications in statistics-simulation and computation* 38.6 (2009), pp. 1228–1234.

[26] Susan Athey and Guido W Imbens. "Identification and inference in nonlinear difference-in-differences models." In: *Econometrica* 74.2 (2006), pp. 431–497.

[27] Andrew Goodman-Bacon. “Difference-in-differences with variation in treatment timing.” In: *Journal of Econometrics* 225.2 (2021), pp. 254–277.

[28] Clifford C Clogg, Eva Petkova, and Adamantios Haritou. “Statistical methods for comparing regression coefficients between models.” In: *American journal of sociology* 100.5 (1995), pp. 1261–1293.

[29] Ron Berman and Ayelet Israeli. “The Value of Descriptive Analytics: Evidence from Online Retailers.” In: *Marketing Science* (2022).

[30] Michael Braun and David A Schweidel. “Modeling customer lifetimes with multiple causes of churn.” In: *Marketing Science* 30.5 (2011), pp. 881–902.

[31] Raghuram Iyengar, Young-Hoon Park, and Qi Yu. “The impact of subscription programs on customer purchases.” In: *Journal of Marketing Research* 59.6 (2022), pp. 1101–1119.

[32] Yiwei Wang, Lauren Xiaoyuan Lu, and Pengcheng Shi. “Does Customer Email Engagement Improve Profitability? Evidence from a Field Experiment in Subscription Service Retailing.” In: *Manufacturing & Service Operations Management* 24.5 (2022), pp. 2703–2721.

[33] Marianne Bertrand, Dean Karlan, Sendhil Mullainathan, Eldar Shafir, and Jonathan Zinman. “What’s advertising content worth? Evidence from a consumer credit marketing field experiment.” In: *The quarterly journal of economics* 125.1 (2010), pp. 263–306.

[34] K Sudhir, Subroto Roy, and Mathew Cherian. “Do sympathy biases induce charitable giving? The effects of advertising content.” In: *Marketing Science* 35.6 (2016), pp. 849–869.

[35] Navdeep S Sahni, S Christian Wheeler, and Pradeep Chintagunta. “Personalization in email marketing: The role of noninformative advertising content.” In: *Marketing Science* 37.2 (2018), pp. 236–258.

[36] Peter E Rossi, Robert E McCulloch, and Greg M Allenby. “The value of purchase history data in target marketing.” In: *Marketing Science* 15.4 (1996), pp. 321–340.

[37] Puneet Manchanda, Jean-Pierre Dubé, Khim Yong Goh, and Pradeep K Chintagunta. “The effect of banner advertising on internet purchasing.” In: *Journal of Marketing Research* 43.1 (2006), pp. 98–108.

[38] Stefan Wager and Susan Athey. “Estimation and inference of heterogeneous treatment effects using random forests.” In: *Journal of the American Statistical Association* 113.523 (2018), pp. 1228–1242.

Appendix A

Appendix for Chapter 1

A.1 Before and After Matching Comparison

We visually demonstrate covariate distribution before (left) and after (right) matching between adopters (blue) and non-adopters (red).

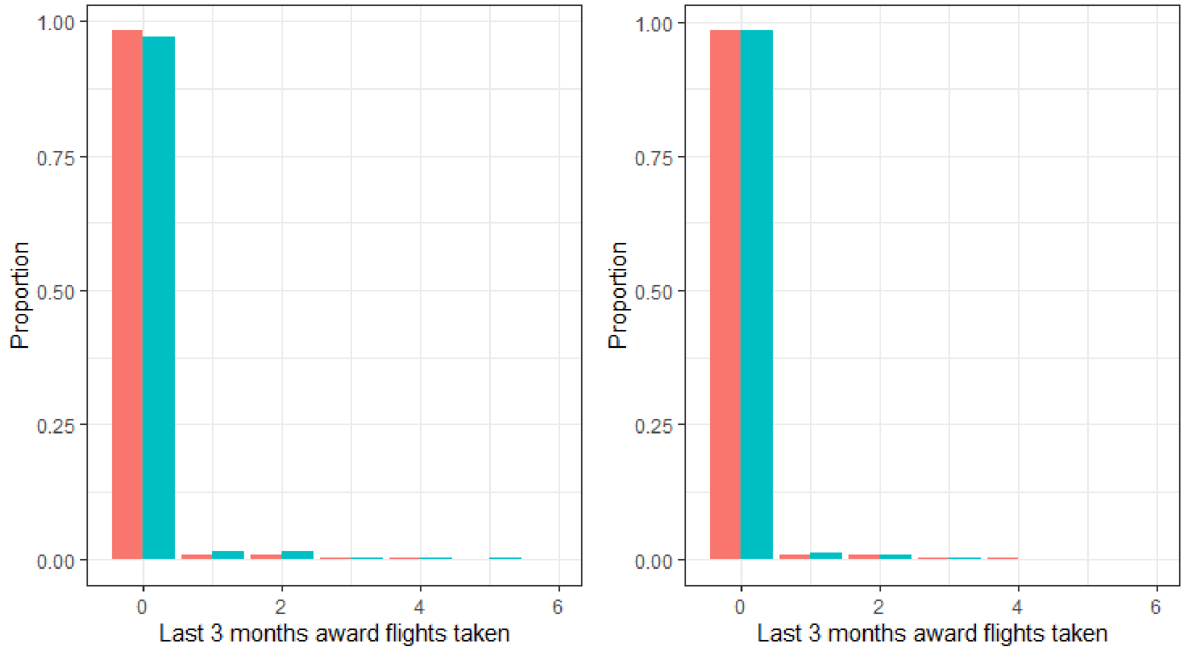

Figure A1: Before and After: Last 3 Months' Number of Award Flights

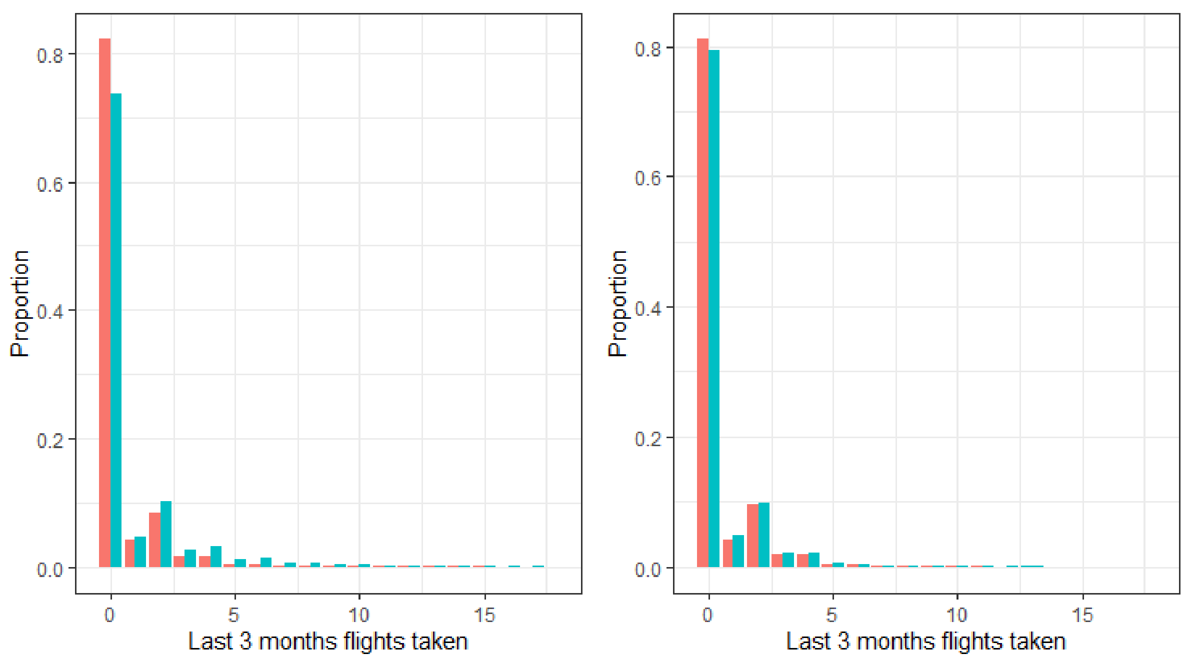

Figure A2: Before and After: Last 3 Months' Number of Flights

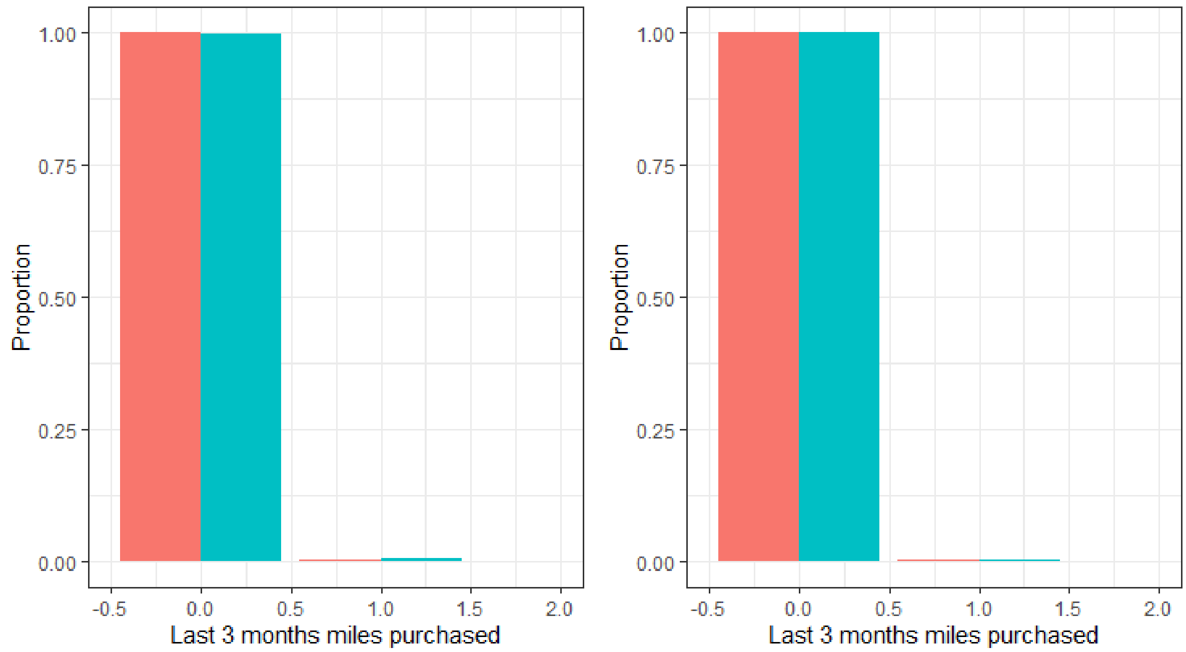

Figure A3: Before and After: Last 3 Months' Miles Purchased

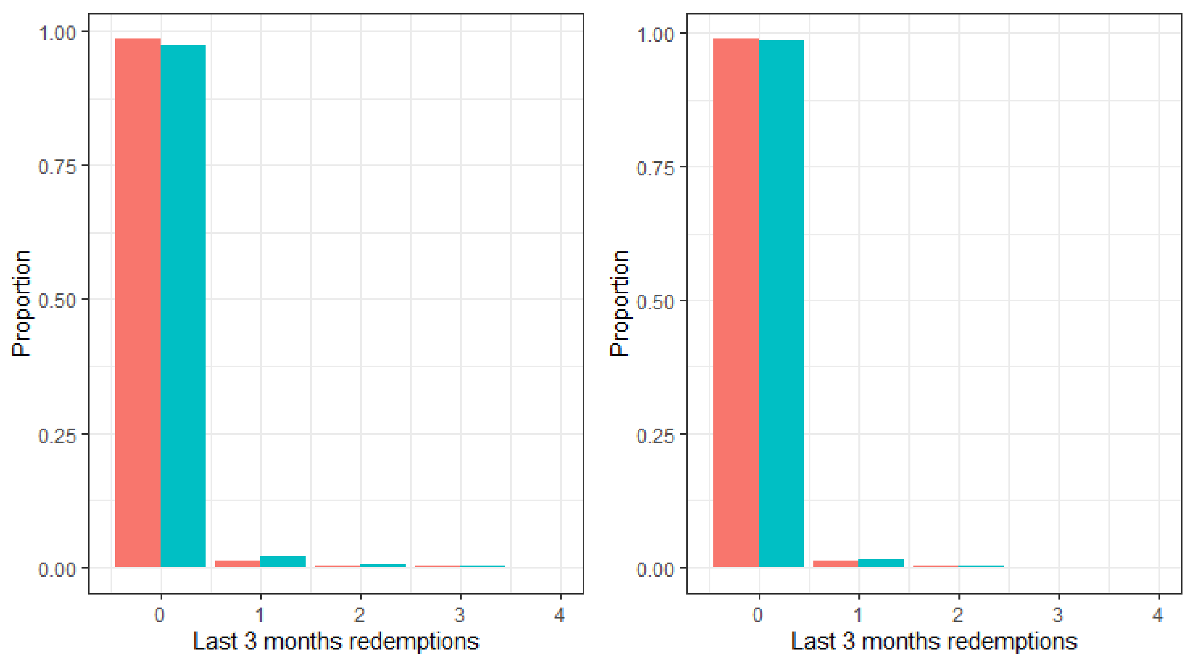

Figure A4: Before and After: Last 3 Months' Number of Redemptions

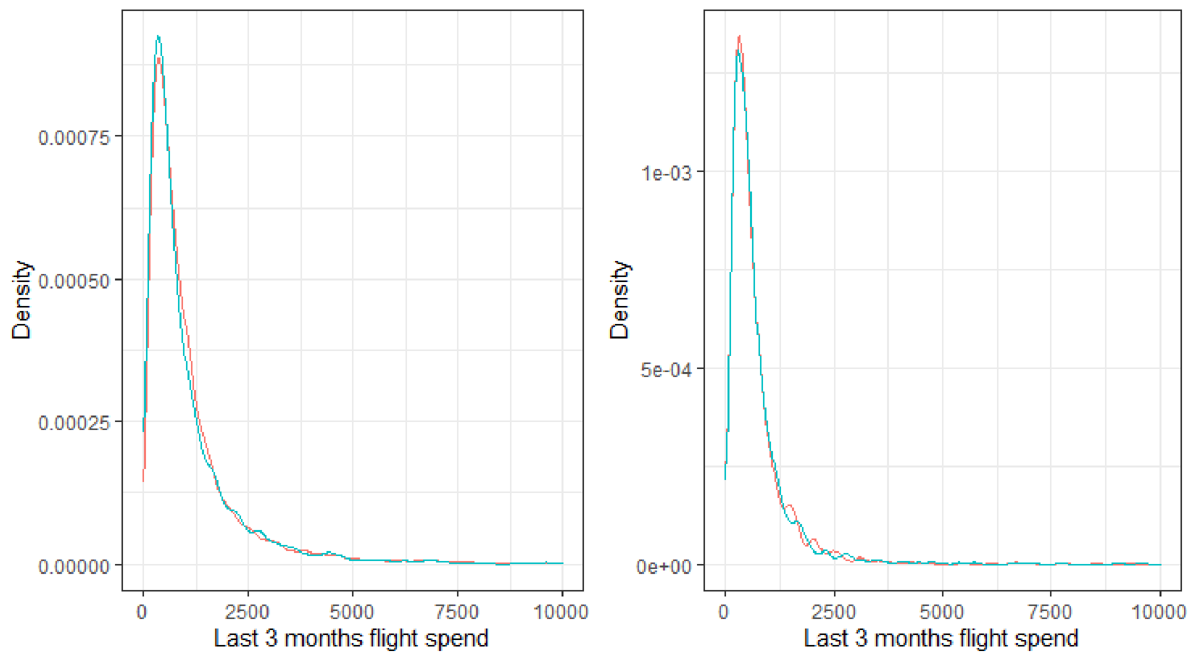

Figure A5: Before and After: Last 3 Months' Flight Spend

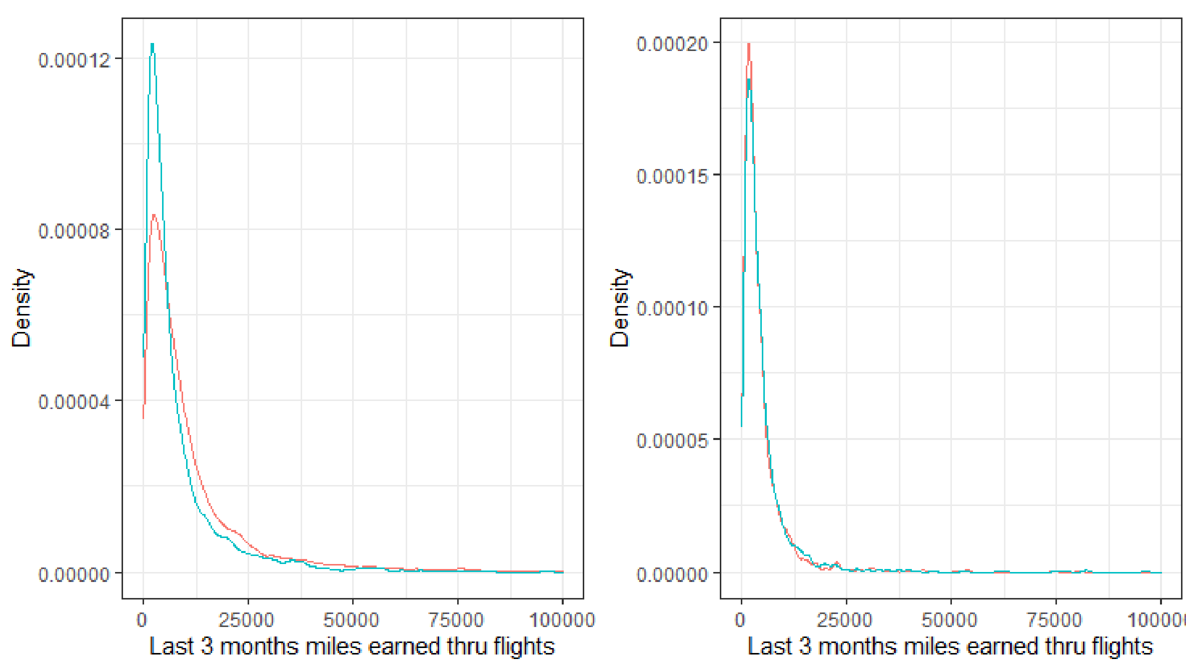

Figure A6: Before and After: Last 3 Months' Miles Earned Through Flights

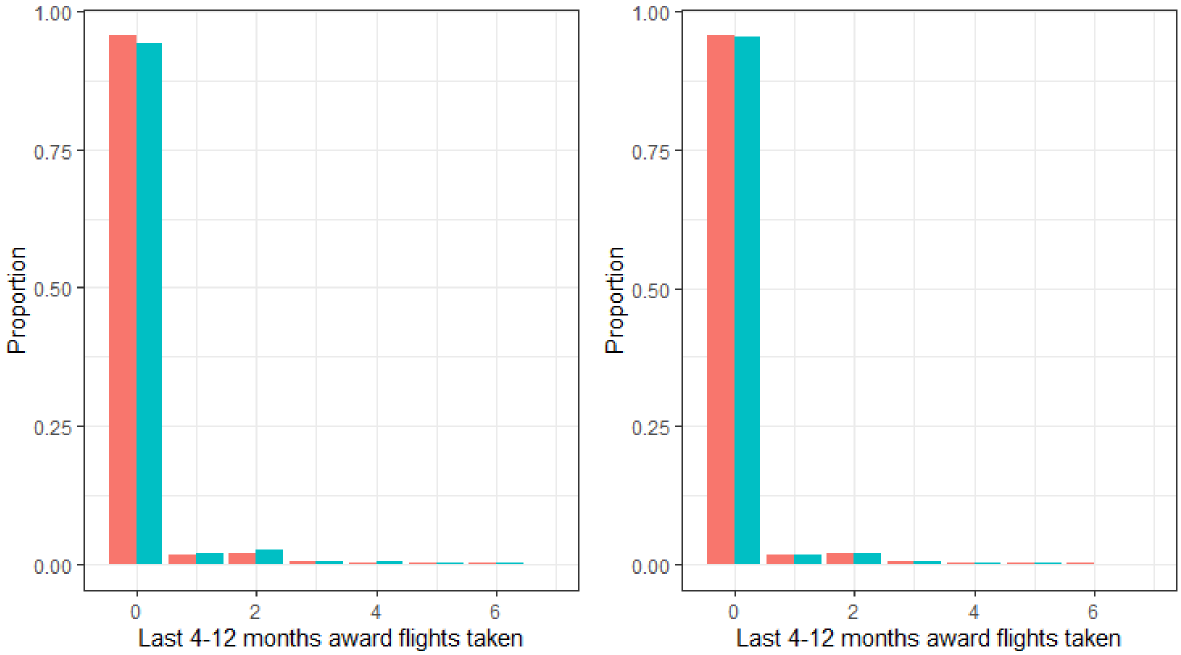

Figure A7: Before and After: Last 4-12 Months' Number of Award Flights

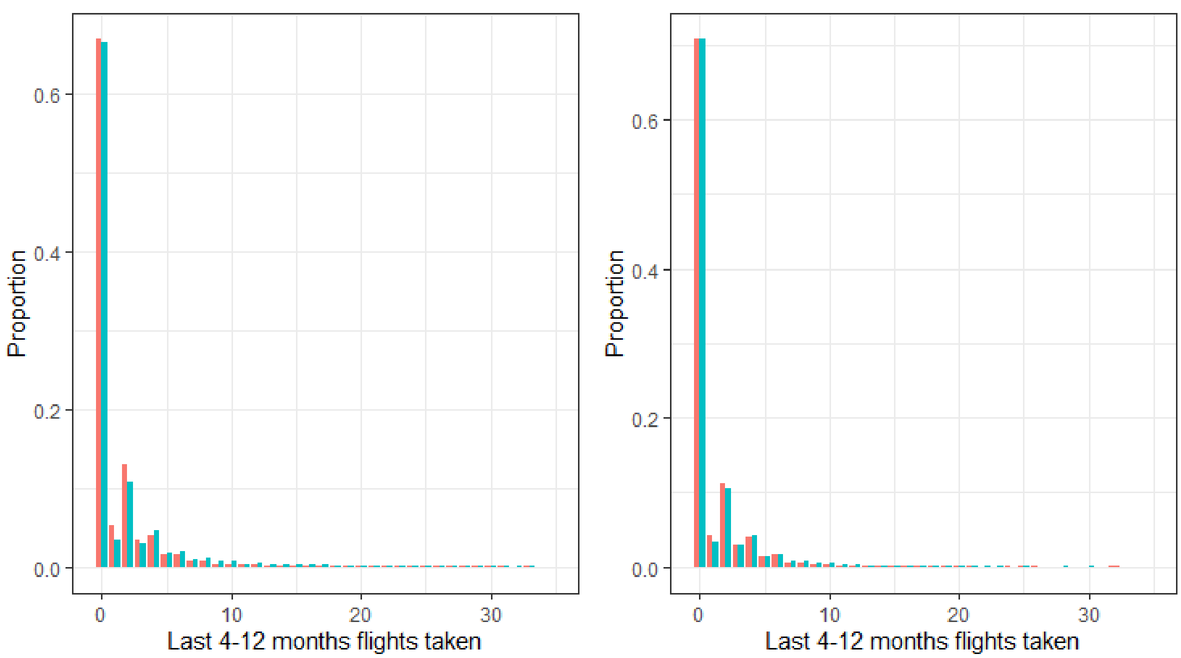

Figure A8: Before and After: Last 4-12 Months' Number of Flights

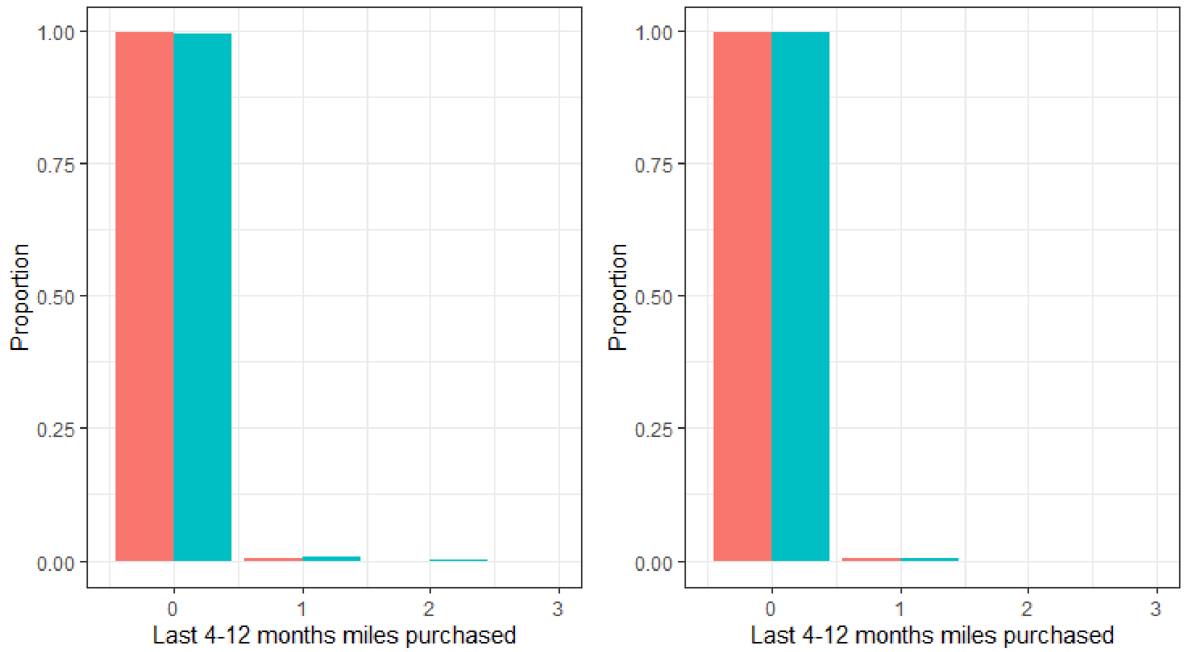

Figure A9: Before and After: Last 4-12 Months' Miles Purchased

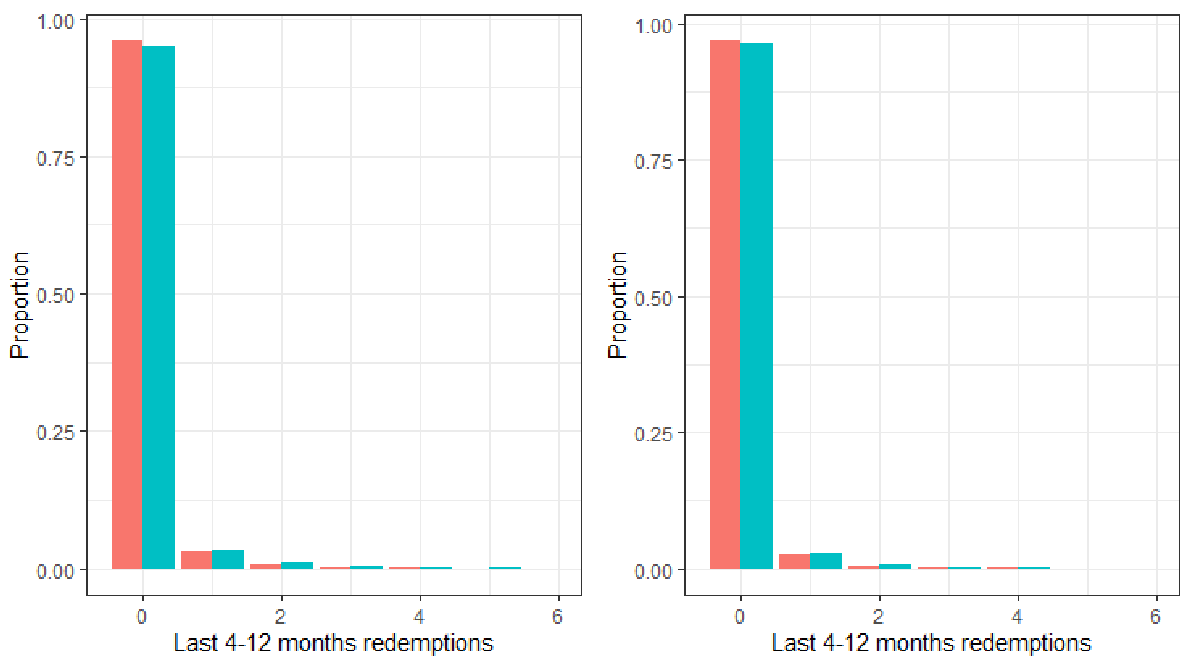

Figure A10: Before and After: Last 4-12 Months' Number of Redemptions

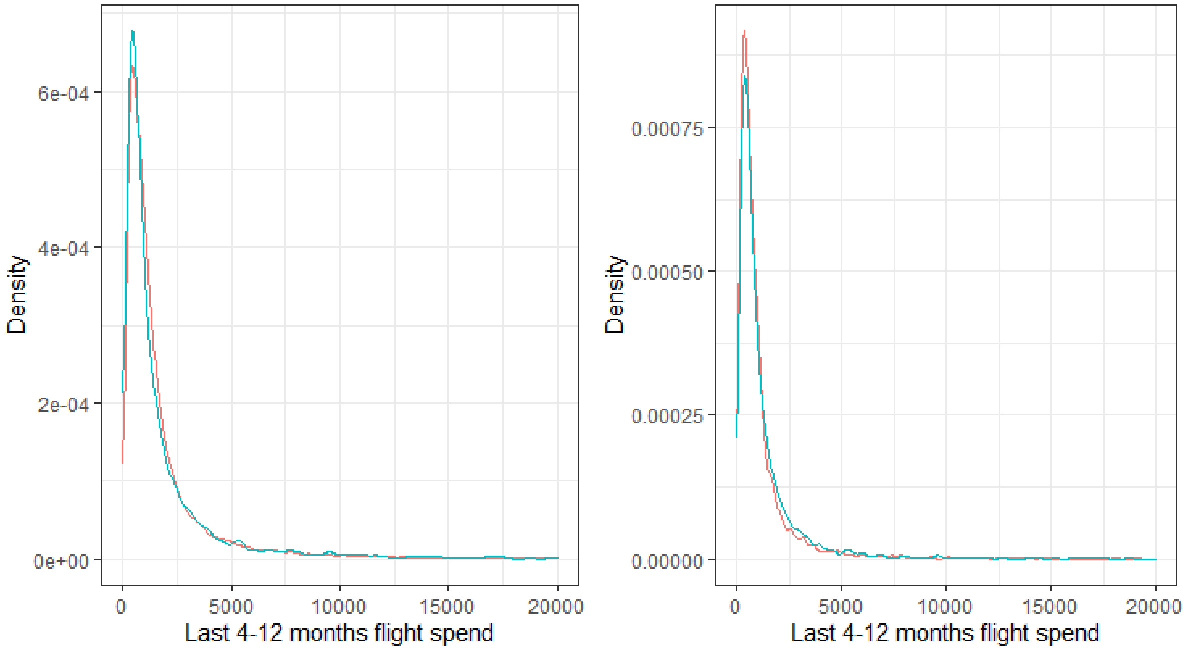

Figure A11: Before and After: Last 4-12 Months' Flight Spend

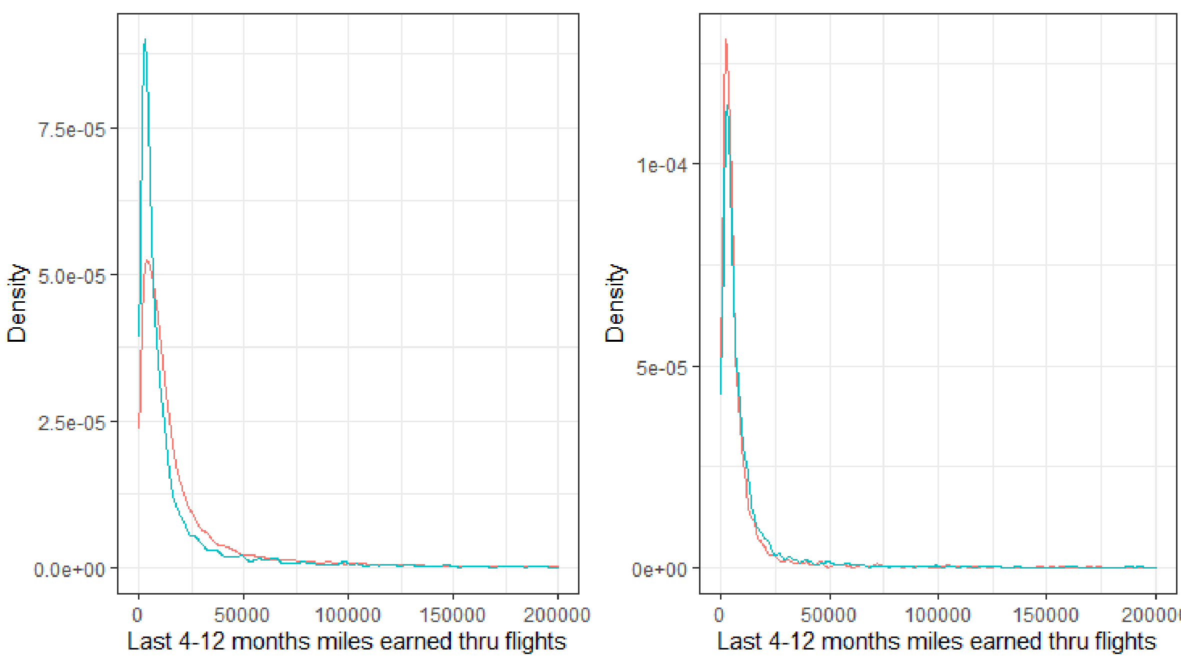

Figure A12: Before and After: Last 4-12 Months' Miles Earned Through Flights

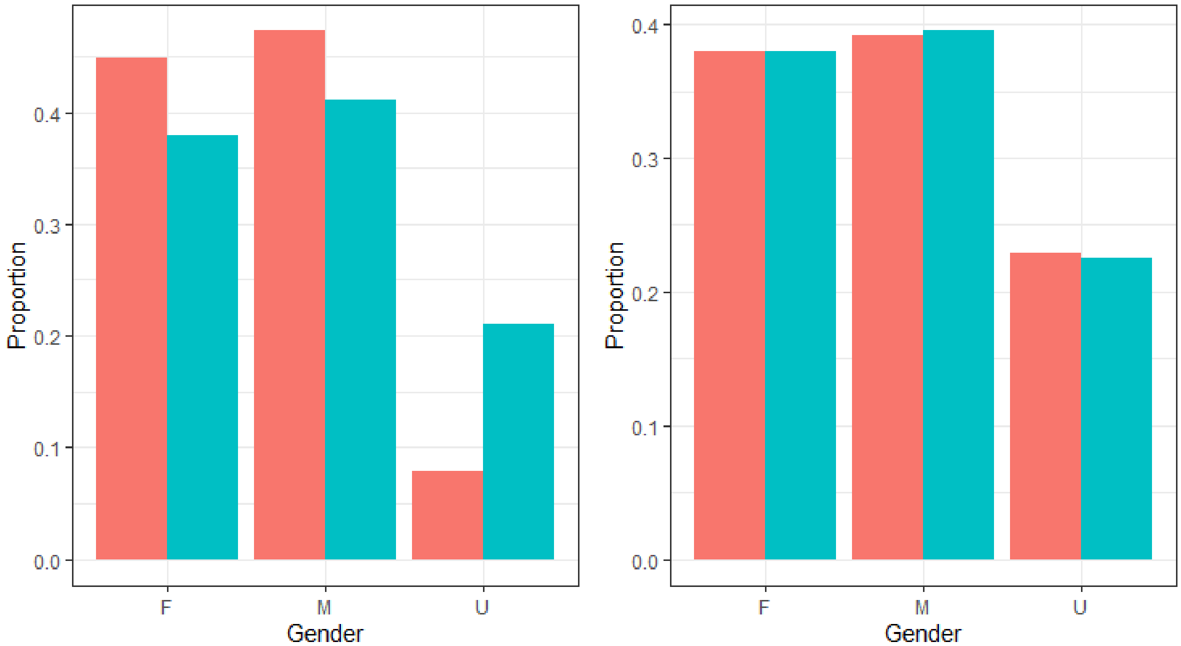

Figure A13: Before and After: Gender

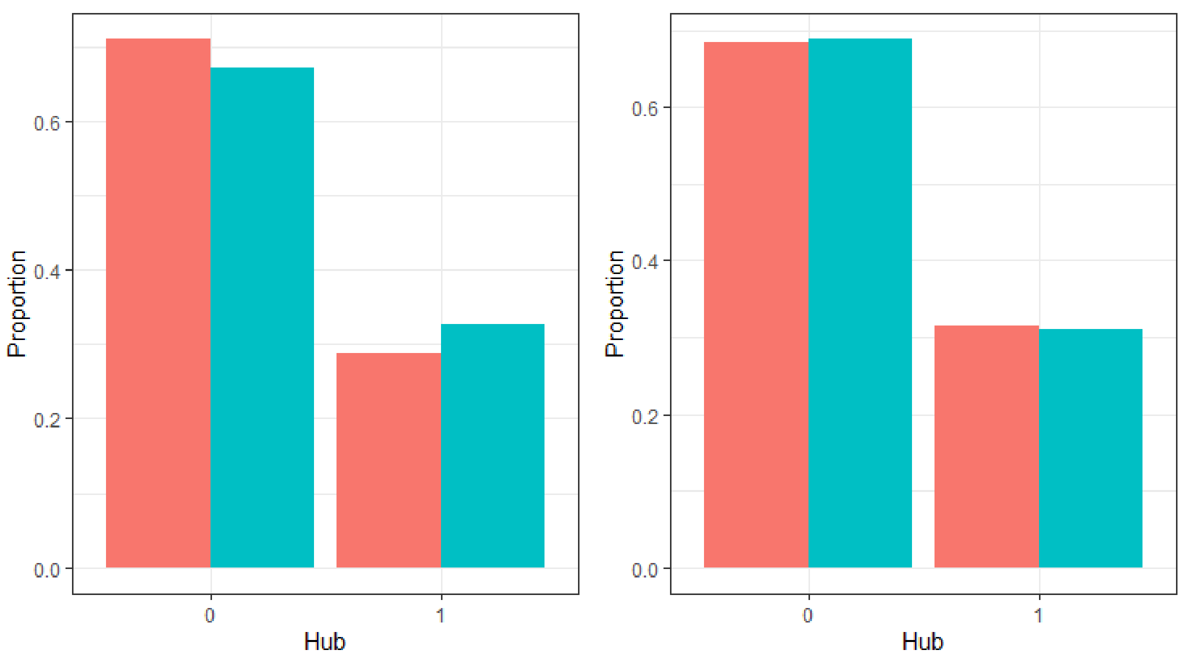

Figure A14: Before and After: Hub

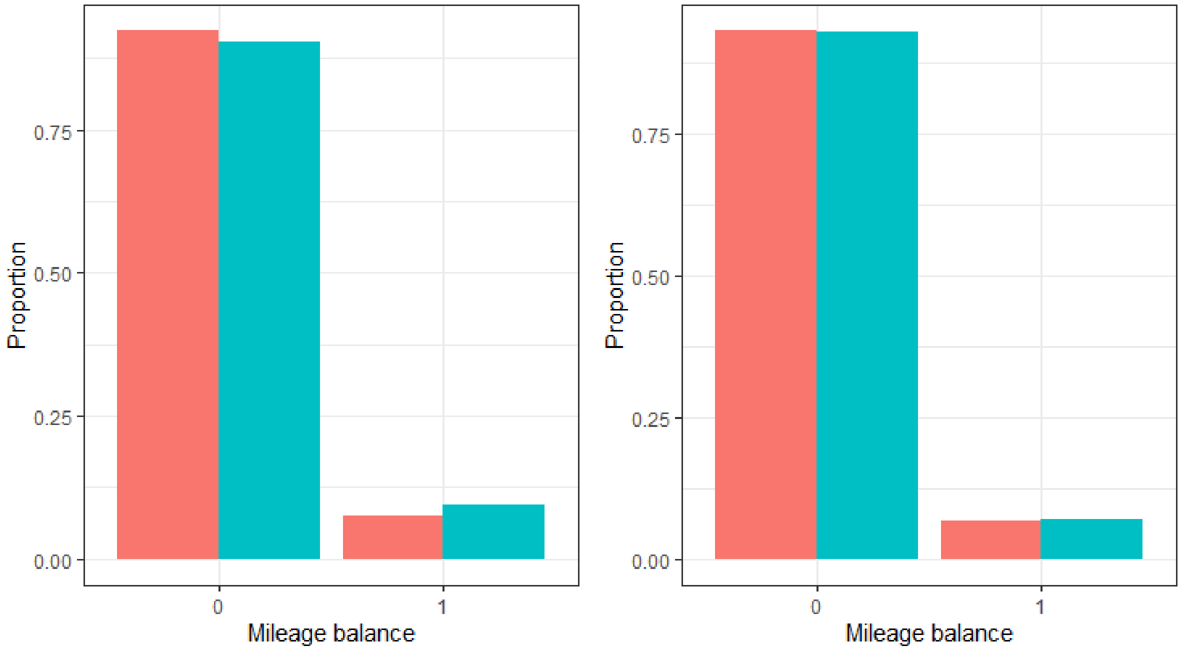

Figure A15: Before and After: Mileage Balance

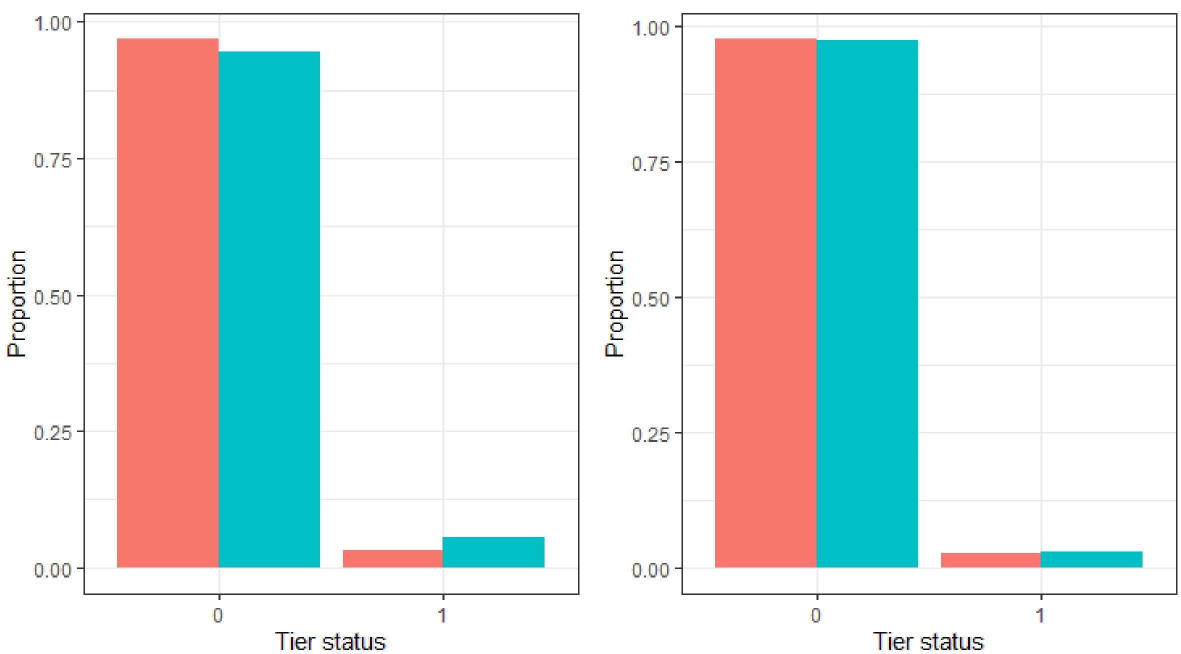

Figure A16: Before and After: Tier Status

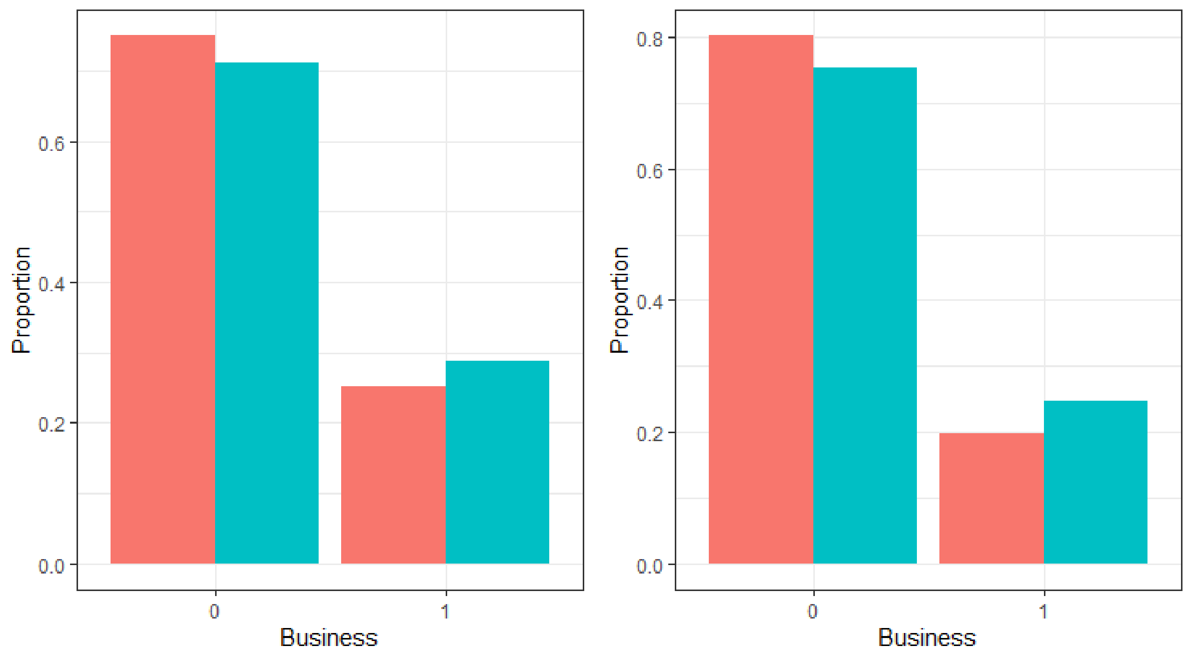

Figure A17: Before and After: Business

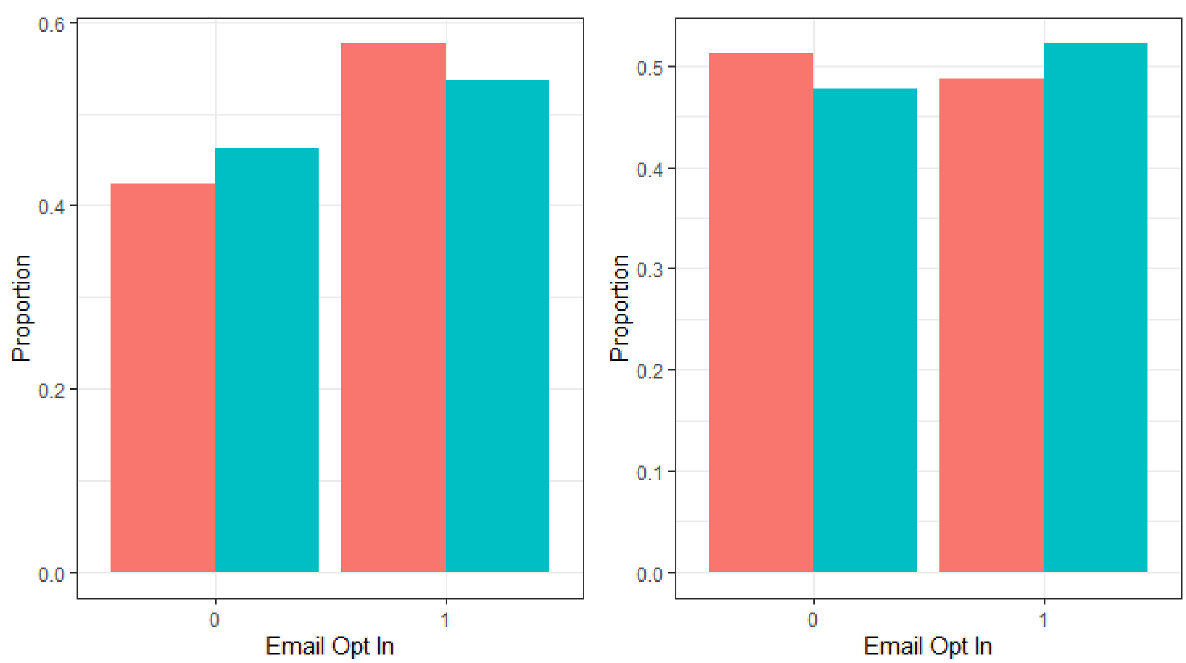

Figure A18: Before and After: Email Opt In

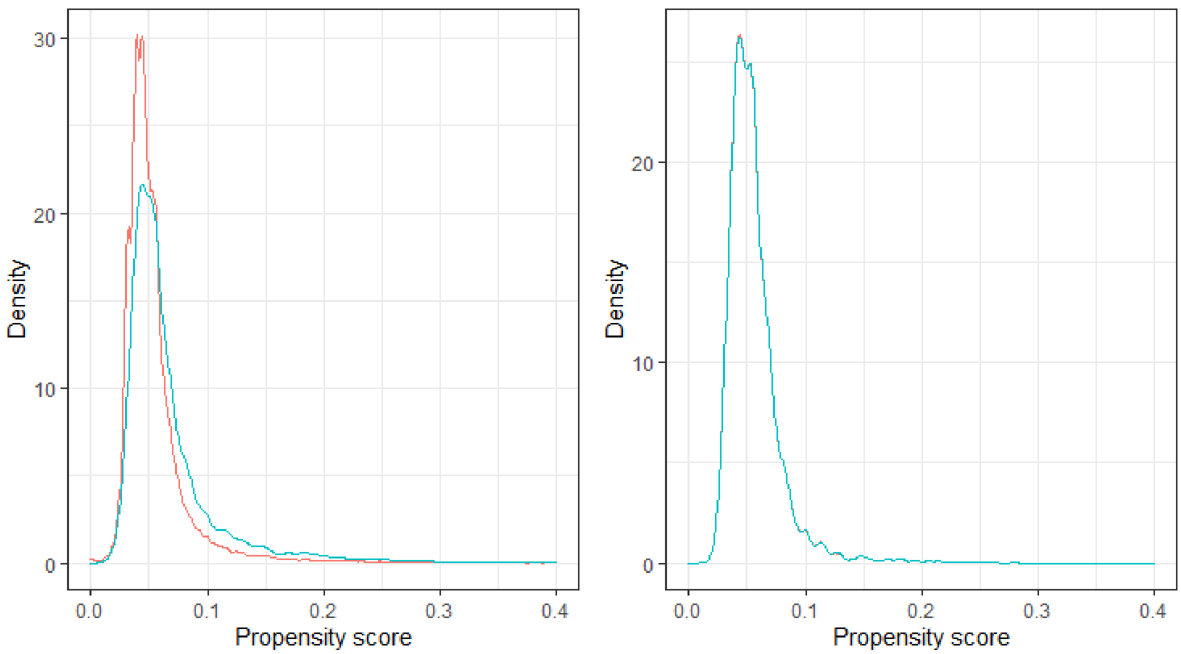

Figure A19: Before and After: Propensity Score

A.2 Placebo Test on the Unmatched Sample

Table A.1: Placebo Tests for Parallel Trend Before Matching

	Dependent variable:				
	Num of Paid Flights	Num of Award Flights	Miles earned thru Flights	Flight Spend	Num of Partner Redemptions
	(1)	(2)	(3)	(4)	(5)
Placebo Adoption	0.058***	0.001	86.987***	14.090***	−0.0001
	(0.004)	(0.001)	(10.458)	(1.341)	(0.0003)
Observations	3,780,264	3,780,264	3,780,264	3,780,264	3,780,264
R^2	0.352	0.149	0.441	0.358	0.132
Individual FE	Y	Y	Y	Y	Y
Year-month FE	Y	Y	Y	Y	Y
Note:Standard errors are clustered at the individual level					*p<0.05; **p<0.01; ***p<0.001

By performing the same placebo test (please see section 4.2) on the unmatched sample, we find that parallel trend assumption is violated for the majority of the DVs, which adds support to why directly using DID yields biased results. The positive and statistically significant coefficients also indicate the gap between adopters and non-adopters widens over time.

A.3 Heterogeneous Treatment Effects on the Treated

Table A.2: Number of Paid Flights

	Segment:										
	All	Hub	Non-hub	High Mileage	Low Mileage	High Bonus	Low Bonus	Peak Season	Off-peak Season	New Members	Existing Members
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Adoption*short	0.239***	0.269***	0.225***	0.273***	0.236***	0.268***	0.207***	0.235***	0.242***	0.193***	0.269***
	(0.005)	(0.010)	(0.006)	(0.026)	(0.005)	(0.008)	(0.006)	(0.007)	(0.007)	(0.006)	(0.007)
Adoption*medium	0.095***	0.122***	0.084***	0.142***	0.092***	0.120***	0.069***	0.095***	0.095***	0.085***	0.102***
	(0.004)	(0.008)	(0.005)	(0.026)	(0.004)	(0.006)	(0.005)	(0.006)	(0.006)	(0.005)	(0.006)
Adoption*long	0.064***	0.098***	0.048***	0.091**	0.062***	0.080***	0.048***	0.067***	0.060***	0.058***	0.067***
	(0.004)	(0.009)	(0.005)	(0.028)	(0.004)	(0.007)	(0.005)	(0.006)	(0.007)	(0.005)	(0.007)
Observations	1,513,824	471,168	1,042,656	105,120	1,408,704	789,600	724,224	783,840	729,984	614,976	898,848
R^2	0.251	0.241	0.256	0.358	0.205	0.262	0.222	0.243	0.259	0.154	0.254
Individual FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year-month FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note:Standard errors are clustered at the individual level *p<0.05; **p<0.01; ***p<0.001

Table A.3: Spend per Flight

	Segment:										
	All	Hub	Non-hub	High Mileage	Low Mileage	High Bonus	Low Bonus	Peak Season	Off-peak Season	New Members	Existing Members
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Adoption*short	−1.383	−7.853	2.833	−10.856	2.864	−4.221	5.685	−7.537	5.388	24.870**	−2.989
	(3.833)	(6.626)	(4.669)	(11.382)	(3.724)	(4.944)	(5.891)	(6.099)	(5.102)	(8.639)	(4.080)
Adoption*medium	9.769*	6.231	11.998*	1.241	14.043**	7.514	15.310*	5.234	14.396*	26.590**	9.792*
	(4.551)	(7.712)	(5.597)	(12.228)	(4.443)	(5.726)	(7.121)	(6.997)	(5.612)	(8.838)	(4.885)
Adoption*long	0.497	−10.053	7.482	−11.794	5.759	0.496	1.503	−2.382	2.362	15.694	0.668
	(5.114)	(8.422)	(6.450)	(13.324)	(5.184)	(6.318)	(8.708)	(7.685)	(6.875)	(10.132)	(5.535)
Observations	163,182	62,019	101,163	33,316	129,866	106,022	57,160	82,152	81,030	25,406	137,776
R^2	0.336	0.301	0.360	0.250	0.375	0.327	0.356	0.330	0.344	0.491	0.313
Individual FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year-month FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note:Standard errors are clustered at the individual level *p<0.05; **p<0.01; ***p<0.001

Table A.4: Miles Earned thru Flights

	Segment:										
	All	Hub	Non-hub	High Mileage	Low Mileage	High Bonus	Low Bonus	Peak Months	Off-peak Months	New Members	Existing Members
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Adoption*short	290.693***	326.568***	274.638***	363.382***	285.045***	360.942***	213.282***	273.862***	306.324***	206.127***	347.888***
	(10.065)	(20.549)	(11.280)	(82.730)	(8.893)	(16.744)	(10.307)	(13.795)	(14.821)	(9.448)	(15.657)
Adoption*medium	156.152***	199.655***	136.724***	269.231**	147.728***	212.998***	92.635***	153.284***	158.476***	122.024***	179.598***
	(9.875)	(19.956)	(11.150)	(84.911)	(8.536)	(16.101)	(10.742)	(14.414)	(13.322)	(10.533)	(15.014)
Adoption*long	117.770***	172.933***	92.235***	148.947	115.519***	158.104***	76.272***	118.532***	112.789***	93.784***	133.439***
	(11.675)	(21.977)	(13.720)	(89.790)	(10.504)	(19.107)	(13.100)	(17.014)	(16.299)	(12.165)	(17.693)
Observations	1,513,824	471,168	1,042,656	105,120	1,408,704	789,600	724,224	783,840	729,984	614,976	898,848
R^2	0.238	0.238	0.238	0.284	0.195	0.247	0.211	0.236	0.241	0.157	0.242
Individual FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year-month FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note:Standard errors are clustered at the individual level *p<0.05; **p<0.01; ***p<0.001

Table A.5: Number of Award flights

	Segment:										
	All	Hub	Non-hub	High Mileage	Low Mileage	High Bonus	Low Bonus	Peak Season	Off-peak Season	New Members	Existing Members
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Adoption*short	0.013***	0.015***	0.012***	0.046***	0.010***	0.018***	0.007***	0.013***	0.012***	0.006***	0.017***
	(0.001)	(0.002)	(0.001)	(0.007)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Adoption*medium	0.039***	0.043***	0.038***	0.054***	0.038***	0.060***	0.016***	0.040***	0.038***	0.029***	0.047***
	(0.001)	(0.002)	(0.001)	(0.005)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Adoption*long	0.040***	0.048***	0.036***	0.059***	0.039***	0.059***	0.020***	0.042***	0.038***	0.030***	0.047***
	(0.001)	(0.002)	(0.001)	(0.005)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	1,513,824	471,168	1,042,656	105,120	1,408,704	789,600	724,224	783,840	729,984	614,976	898,848
R^2	0.058	0.062	0.055	0.065	0.055	0.059	0.055	0.058	0.058	0.045	0.059
Individual FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year-month FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note:Standard errors are clustered at the individual level *p<0.05; **p<0.01; ***p<0.001

Table A.6: Number of Partner Redemptions

	Segment:										
	All	Hub	Non-hub	High Mileage	Low Mileage	High Bonus	Low Bonus	Peak Season	Off-peak Season	New Members	Existing Members
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Adoption*short	0.003***	0.003***	0.004***	0.004	0.003***	0.004***	0.002***	0.004***	0.003***	0.003***	0.004***
	(0.0004)	(0.001)	(0.001)	(0.002)	(0.0004)	(0.001)	(0.0004)	(0.001)	(0.001)	(0.0004)	(0.001)
Adoption*medium	0.006***	0.006***	0.006***	0.006***	0.006***	0.009***	0.003***	0.007***	0.005***	0.007***	0.005***
	(0.0004)	(0.001)	(0.001)	(0.002)	(0.0004)	(0.001)	(0.0003)	(0.001)	(0.0004)	(0.001)	(0.0004)
Adoption*long	0.006***	0.006***	0.006***	0.003	0.006***	0.007***	0.004***	0.006***	0.006***	0.008***	0.005***
	(0.0004)	(0.001)	(0.0005)	(0.002)	(0.0004)	(0.001)	(0.0004)	(0.001)	(0.0005)	(0.001)	(0.0005)
Observations	1,513,824	471,168	1,042,656	105,120	1,408,704	789,600	724,224	783,840	729,984	614,976	898,848
R^2	0.052	0.038	0.058	0.050	0.052	0.057	0.039	0.059	0.044	0.066	0.046
Individual FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year-month FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note:Standard errors are clustered at the individual level *p<0.05; **p<0.01; ***p<0.001

A.4 P-values of the Differences Between HTT and ATT

Table A.7: Flight Spend

	Segment:									
	Hub	Non-hub	High Mileage	Low Mileage	High Bonus	Low Bonus	Peak Season	Off-peak Season	New Members	Existing Members
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Adoption*short	0.335	0.538	0.857	0.940	0.000	0.000	0.613	0.659	0.000	0.003
Adoption*medium	0.042	0.173	0.147	0.535	0.001	0.000	0.968	0.982	0.013	0.174
Adoption*long	0.017	0.103	0.589	0.828	0.030	0.006	0.705	0.540	0.096	0.400

Table A.8: Number of Paid Flights

	Segment:									
	Hub	Non-hub	High Mileage	Low Mileage	High Bonus	Low Bonus	Peak Season	Off-peak Season	New Members	Existing Members
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Adoption*short	0.007	0.073	0.199	0.671	0.002	0.000	0.642	0.727	0.000	0.000
Adoption*medium	0.003	0.086	0.074	0.596	0.001	0.000	1.000	1.000	0.118	0.332
Adoption*long	0.001	0.012	0.340	0.724	0.047	0.012	0.677	0.620	0.349	0.710

Table A.9: Spend per Flight

	Segment:									
	Hub	Non-hub	High Mileage	Low Mileage	High Bonus	Low Bonus	Peak Season	Off-peak Season	New Members	Existing Members
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Adoption*short	0.398	0.485	0.430	0.427	0.650	0.315	0.393	0.289	0.005	0.774
Adoption*medium	0.693	0.757	0.513	0.502	0.758	0.512	0.587	0.522	0.091	0.997
Adoption*long	0.284	0.396	0.389	0.470	1.000	0.921	0.755	0.828	0.181	0.982

Table A.10: Miles Earned thru Flights

	Segment:									
	Hub	Non-hub	High Mileage	Low Mileage	High Bonus	Low Bonus	Peak Season	Off-peak Season	New Members	Existing Members
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Adoption*short	0.117	0.288	0.383	0.674	0.000	0.000	0.324	0.383	0.000	0.002
Adoption*medium	0.051	0.192	0.186	0.519	0.003	0.000	0.870	0.889	0.018	0.192
Adoption*long	0.027	0.156	0.731	0.886	0.072	0.018	0.971	0.804	0.155	0.460

Table A.11: Number of Award Flights

	Segment:									
	Hub	Non-hub	High Mileage	Low Mileage	High Bonus	Low Bonus	Peak Season	Off-peak Season	New Members	Existing Members
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Adoption*short	0.371	0.480	0.000	0.034	0.025	0.000	1.000	0.480	0.000	0.074
Adoption*medium	0.074	0.480	0.003	0.480	0.000	0.000	0.480	0.480	0.000	0.000
Adoption*long	0.000	0.005	0.000	0.480	0.000	0.000	0.157	0.157	0.000	0.000

Table A.12: Number of Partner Redemptions

	Segment:									
	Hub	Non-hub	High Mileage	Low Mileage	High Bonus	Low Bonus	Peak Season	Off-peak Season	New Members	Existing Members
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Adoption*short	1.000	0.353	0.624	1.000	0.353	0.077	0.353	1.000	1.000	0.353
Adoption*medium	1.000	1.000	1.000	1.000	0.005	0.000	0.353	0.077	0.353	0.077
Adoption*long	1.000	1.000	0.141	1.000	0.353	0.000	1.000	1.000	0.063	0.118

A.5 P-values of the Differences Between Pair-wise Segments

Table A.13: Flight Spend

	Segment:				
	Hub vs. Non-hub (1)	High vs. Low Mileage (2)	High vs. Low Bonus (3)	Peak vs. Off-peak Season (4)	New vs. Existing Members (5)
Adoption*short	0.178	0.844	0.000	0.413	0.000
Adoption*medium	0.004	0.119	0.000	0.987	0.001
Adoption*long	0.001	0.559	0.000	0.393	0.032

Table A.14: Number of Paid Flights

	Segment:				
	Hub vs. Non-hub (1)	High vs. Low Mileage (2)	High vs. Low Bonus (3)	Peak vs. Off-peak Season (4)	New vs. Existing Members (5)
Adoption*short	0.000	0.162	0.000	0.480	0.000
Adoption*medium	0.000	0.057	0.000	1.000	0.030
Adoption*long	0.000	0.305	0.000	0.448	0.295

Table A.15: Spend per Flight

	Segment:				
	Hub vs. Non-hub (1)	High vs. Low Mileage (2)	High vs. Low Bonus (3)	Peak vs. Off-peak Season (4)	New vs. Existing Members (5)
Adoption*short	0.187	0.252	0.198	0.104	0.004
Adoption*medium	0.545	0.325	0.394	0.307	0.096
Adoption*long	0.098	0.220	0.925	0.645	0.193

Table A.16: Miles Earned thru Flights

	Segment:				
	Hub vs. Non-hub (1)	High vs. Low Mileage (2)	High vs. Low Bonus (3)	Peak vs. Off-peak Season (4)	New vs. Existing Members (5)
Adoption*short	0.027	0.346	0.000	0.109	0.000
Adoption*medium	0.006	0.155	0.000	0.791	0.002
Adoption*long	0.002	0.712	0.000	0.807	0.065

Table A.17: Number of Award Flights

	Segment:				
	Hub vs. Non-hub (1)	High vs. Low Mileage (2)	High vs. Low Bonus (3)	Peak vs. Off-peak Season (4)	New vs. Existing Members (5)
Adoption*short	0.180	0.000	0.000	0.480	0.000
Adoption*medium	0.025	0.002	0.000	0.157	0.000
Adoption*long	0.000	0.000	0.000	0.005	0.000

Table A.18: Number of Partner Redemptions

	Segment:				
	Hub vs. Non-hub	High vs. Low Mileage	High vs. Low Bonus	Peak vs. Off-peak Season	New vs. Existing Members
	(1)	(2)	(3)	(4)	(5)
Adoption*short	0.480	0.624	0.063	0.480	0.353
Adoption*medium	1.000	1.000	0.000	0.063	0.063
Adoption*long	1.000	0.141	0.005	1.000	0.007

Appendix B

Appendix for Chapter 2

Table B.1: First, Second and Third Order Rates by Email Condition

	Control	Convenience	Price	Reminder	Risk	Safety
Number of Customers	418,925	1,502,035	1,501,944	1,502,090	1,502,084	1,501,912
First Order Rate	0.2561%	0.2717%	0.2626%	0.2750%	0.2690%	0.2675%
	(0.0078%)	(0.0043%)	(0.0043%)	(0.0042%)	(0.0042%)	(0.0042%)
	[0.48]	[0.13]		[0.04]	[0.29]	[0.41]
Second Order Rate	0.1590%	0.1682%	0.1589%	0.1657%	0.1677%	0.1623%
	(0.0062%)	(0.0033%)	(0.0033%)	(0.0033%)	(0.0033%)	(0.0033%)
	[1]	[0.05]		[0.14]	[0.06]	[0.46]
Third Order Rate	0.1050%	0.1120%	0.1019%	0.1087%	0.1101%	0.1085%
	(0.0050%)	(0.0027%)	(0.0027%)	(0.0027%)	(0.0027%)	(0.0026%)
	[0.60]	[0.01]		[0.07]	[0.03]	[0.08]

Note: Standard Errors in Parentheses and P-values vs. Price in Brackets

Table B.2: Overall Cancellation Rate - Subscription

	Control	Convenience	Price	Reminder	Risk	Safety
Number of Subscriptions	1,746	6,426	6,354	6,625	6,407	6,644
Overall Cancellation Rate	72.57%	73.64%	72.87%	73.27%	73.53%	72.62%
	(1.07%)	(0.55%)	(0.56%)	(0.54%)	(0.55%)	(0.55%)
	[0.83]	[0.33]		[0.62]	[0.41]	[0.77]

Note: Standard Errors in Parentheses and P-values vs. Price in Brackets

Table B.3: Conditional Cancellation Rate - Subscription

	Control	Convenience	Price	Reminder	Risk	Safety
Number of Cancelled Subscriptions	1,267	4,732	4,630	4,854	4,711	4,825
Conditional Cancellation Rate within One Order	51.70%	49.73%	51.71%	51.63%	50.01%	52.46%
	(1.16%)	(0.60%)	(0.61%)	(0.60%)	(0.60%)	(0.60%)
	[1]	[0.06]		[0.96]	[0.11]	[0.48]
Conditional Cancellation Rate within Two Orders	76.56%	75.34%	77.67%	75.73%	75.65%	77.02%
	(0.83%)	(0.44%)	(0.45%)	(0.42%)	(0.45%)	(0.43%)
	[0.42]	[0.01]		[0.03]	[0.02]	[0.46]

Note: Standard Errors in Parentheses and P-values vs. Price in Brackets

Table B.4: Summary Statistics under Optimal Assignment

Optimal message	Control	Convenience	Price	Reminder	Risk	Safety
Ecom order 4+	0.0832	0.0635	0.0377	0.0544	0.0559	0.0624
BM order 4+	0.5749	0.4970	0.4537	0.4707	0.5022	0.4804
Ecom purchase irregularity	0.0603	0.0473	0.0287	0.0387	0.0437	0.0468
BM purchase irregularity	0.5118	0.4457	0.3913	0.4174	0.4546	0.4157
Ecom order %	0.1232	0.1030	0.0937	0.1041	0.1082	0.1060
Ecom sale %	0.3379	0.2834	0.2599	0.2824	0.2936	0.2738
BM sale %	0.5549	0.5075	0.4950	0.4746	0.4875	0.4933
Ecom private label %	0.0947	0.0902	0.0758	0.0914	0.0943	0.0785
BM private label %	0.2807	0.3036	0.2901	0.3114	0.2801	0.3125
Log days since last BM purchase	4.0113	4.2444	4.3322	4.3288	4.3606	4.3319
Log days since last Ecom purchase	1.7679	1.4614	1.4042	1.4880	1.5831	1.4057
Previous subs customer	0.0547	0.0388	0.0322	0.0484	0.0392	0.0382