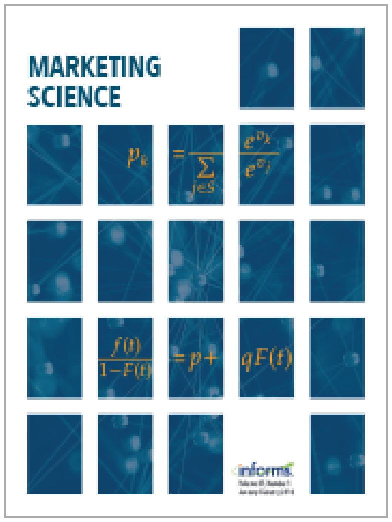


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Can Non-tiered Customer Loyalty Programs Be Profitable?

Arun Gopalakrishnan , Zhenling Jiang , Yulia Nevskaya , Raphael Thomadsen

informs
http://pubsonline.informs.org/journal/mksc

MARKETING SCIENCE
Vol. 40, No. 3, May-June 2021, pp. 508–526
ISSN 0732-2399 (print), ISSN 1526-548X (online)

Can Non-tiered Customer Loyalty Programs Be Profitable?

Arun Gopalakrishnan,[a] Zhenling Jiang,[b] Yulia Nevskaya,[c] Raphael Thomadsen[c]

[a] Jones Graduate School of Business, Rice University, Houston, Texas 77005; [b] Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104; [c] Olin Business School, Washington University in St. Louis, St. Louis, Missouri 63130

Contact: agopala@rice.edu, http://orcid.org/0000-0002-9222-3943 (AG); zhenling@wharton.upenn.edu, http://orcid.org/0000-0002-4581-0125 (ZJ); yulia.nevskaya@wustl.edu, http://orcid.org/0000-0001-5975-6919 (YN); thomadsen@wustl.edu, https://orcid.org/0000-0003-3958-5416 (RT)

Received: February 29, 2016
Revised: June 22, 2017; May 6, 2019; June 19, 2020
Accepted: August 2, 2020
Published Online in Articles in Advance: March 18, 2021

https://doi.org/10.1287/mksc.2020.1268

Abstract. We study the impact of launching a non-tiered customer loyalty program on consumers' spending per visit, frequency of visits and attrition rates, and overall customer value. We demonstrate these results both through descriptive difference-in-difference regressions and a duration-dependent hidden Markov model we develop. We find that the program increases customer value by almost 30% over a five-year horizon, which is considerably larger than has been previously found for non-tiered loyalty programs. Most of the impact of the loyalty program comes through attrition: we show that the program's reduction in attrition accounts for more than 80% of the program's total lift, whereas increased frequency accounts for less than 20% of the program's lift. The program's lift is highest for least and most frequent automatic members, who experience reductions in attrition rates after joining the program. The impact of the loyalty program on spending per visit is negligible.

History: K. Sudhir served as the senior editor and Anja Lambrecht served as associate editor for this article.
Supplemental Material: The data replication files and online appendix are available at https://doi.org/10.1287/mksc.2020.1268.

Keywords: customer loyalty • customer lifetime value • customer relationship management • reward programs

1. Introduction

Customer loyalty programs are used by a wide range of businesses. In 2013, the average U.S. household belonged to 21.9 loyalty programs and actively participated in 9.5 of them (Berry 2013). An important attribute of a loyalty program is whether it provides a tiered (i.e., increased rewards for reaching higher thresholds) or non-tiered reward structure. Hotel chains and airlines typically offer tiered programs, which create economic lock-in because of increasing benefits and consumer self-signaling (Drèze and Nunes 2009, Orhun and Guo 2019). By contrast, non-tiered programs, such as "buy 10 get 1 free" and "$\$X$ off for every $\$Y$ of spend," are popularly used in retail and service industries such as grocery stores, coffee shops, sandwich shops, and golf courses. Unlike a tiered program, it is less obvious how a non-tiered program may increase customer demand or be profitable because status cannot be earned. Furthermore, a non-tiered program provides much weaker economic lock-in from skipping a visit to the particular business because although the customer receives no credit for the one skipped visit, he or she incurs no loss of the increasing value for future visits. However, consumers may respond to loyalty programs for psychological reasons. For example, the presence of the loyalty program can make the customer feel more connected to a particular firm, which can lead to the customer visiting the firm more often.

The existing literature has typically found small or statistically insignificant effects from non-tiered loyalty programs. For example, Sharp and Sharp (1997) find only a weak loyalty-program impact through repeat purchases in retail outlets, and the effect is not consistently observed for all brands. Hartmann and Viard (2008) find that a "buy 10 get 1 free" loyalty program does not create significant switching costs for members. Lewis (2004) finds a 2% revenue increase from a frequency loyalty program. The non-tiered loyalty program Leenheer et al. (2007) study was found to increase a store's share of wallet by about 4%.[1]

One difficulty in measuring the impact of a loyalty program is that each customer cannot be simultaneously observed while in and out of the loyalty program. Measuring the program effect on any behavioral aspect is then contingent on the construction of a plausible counterfactual. The issue would be easily resolved if customers could be randomly assigned to a treatment (*join the program*) and a control (*do not join the program*) group. However, in practice, customers self-select into a program once a loyalty program is launched, leading to large selection effects (van Heerde and Bijmolt 2005, Leenheer et al. 2007). Studies that do not fully account for selection effects (Bolton et al. 2000, Lal and Bell 2003) have shown sizable effects, which are likely to be upwardly biased. Leenheer et al. (2007) address such selection issues by using instrumental variables, but such a solution depends on having strong and valid instruments, which are hard to find. Other papers consider only the behavior of loyalty-program members (Lewis 2004, Liu 2007, Hartmann and Viard 2008, Kopalle et al. 2012, Stourm et al. 2015) and use a model based on rational economic behavior to identify the program's impact. For example, the primary economic rationale for a non-tiered loyalty program would be reward redemption, which then suggests that high-frequency customers would mainly benefit, whereas low-frequency customers would see less value in being part of the program. This approach may not capture the entire impact of the program because a consumer's responsiveness to a loyalty program can go beyond those based on pure economic utility. Other psychological benefits, such as habit or status, may accrue simply from being a member of a loyalty program even if one does not qualify for its rewards (see Henderson et al. 2011 for a review). We allow for this possibility in our analysis.

Our empirical analysis uses data from a men's hair salon chain that introduced a non-tiered loyalty program during the data-collection period. The loyalty program is free to join—customers only need to provide an email address to enroll. Members accrue spending after joining the program and receive a \$5-off reward coupon via email for every \$100 they spend on hair services and hair-care products.

The goal of this paper is to measure the effectiveness of this non-tiered loyalty program, represented by the change in the time-discounted customer value over a five-year horizon. Our main contribution is to demonstrate, with our unique data set combined with a careful modeling approach, that a simple non-tiered loyalty program can be beneficial to a firm and that this benefit is driven mostly by an increase in customer retention. Our analysis shows a 29.5% increase in customer lifetime value (CLV) from a non-tiered loyalty program, which is much larger than the effect previously found in the literature.[2]

Most of the increase in CLV comes from reduced customer attrition. Specifically, we find that the program reduces the probability that customers will enter a long hiatus from patronizing the hair salon (while getting haircuts elsewhere). Although the impact of loyalty programs on attrition has not been studied much in non-contractual settings, the impact of the program on attrition alone can explain 23.6% of the 29.5% observed increased in CLV. Consistent with the prior literature, we find that the loyalty program has a minimal impact on spending per visit and a small impact on visit frequency, which leads to a 4.1% increase in CLV. We also find that the largest percentage increases in CLV occur in the segments with the most and least pre-program visits. Consistent with the overall effect, we find that the impact on the most and least frequent groups is primarily driven by increased retention.

We use a number of novel aspects in our data and estimation approach. First, we collect data for a cohort of more than 5,500 customers who were acquired by the firm in the same quarterly period in order to ensure that these customers are comparable in terms of the amount of time they have to become acquainted with the firm. All subsequent-visit data from these consumers are captured for 30 months, during which time the loyalty program was introduced. Observing consumer behavior before and after the program's introduction allows us to control for consumer heterogeneity and selection effects. Second, we exploit an institutional detail in our setting: some customers were automatically enrolled in the loyalty program at program introduction such that their timing of joining the program is exogenous. We further divide customers into segments based on their pre-program visit frequency and use this metric to create matched segments of automatically enrolled customers (which we hereafter refer to as *automatic members*) and non-enrolled customers[3] (which we refer to as *non-members*).

On the methodology side, we bring together the evaluation of non-tiered loyalty programs and the literature on CLV models (Schmittlein et al. 1987, Fader et al. 2010, Ascarza and Hardie 2013). Specifically, we develop a modeling framework based on a hidden Markov model (HMM; Netzer et al. 2008) that integrates customer attrition and frequency of visits. This model allows for estimation of the program's overall effectiveness and evaluation of the program's impact on customer visit frequency and attrition separately. By estimating separate HMM-based models for each of these segments (based on the number of pre-program visits) and customer types (automatic versus non-members), we allow for a flexible accounting of heterogeneity by avoiding pooling data from infrequent and frequent customers. We compute the changes in behavior after program introduction for automatic members and non-members and attribute the difference in the change between automatic and non-members to the effect of the program (i.e., a difference-in-differences approach).

The rest of this paper is organized as follows. In Section 2, we describe our data and the structure of the loyalty program and lay out some descriptive analyses. In Section 3, we describe the model specification of the HMM-based approach we use for analysis. In Section 4, we discuss the estimated model and analyze the impact of the loyalty program on customer value. We present our conclusions in Section 5.

2. Data Set and Descriptive Analysis

In this section, we first describe our data in Section 2.1. We then provide some descriptive analyses in Section 2.2 that demonstrate the impact of the loyalty program as a precursor to an HMM-based approach that we introduce later in Section 3.

2.1. Data Set

The empirical analysis is based on a data set obtained from a chain of men's hair salons. We observe a cohort of customers acquired between six and nine months before the launch of a loyalty program.[4] All visits from these customers are captured for a period of 30 months. In the 10th month (of the 30-month period), the company introduced a customer loyalty program. Therefore, we have observations of customers' visit behavior both before and after the launch of the loyalty program. Each transaction record contains a unique customer identification number, the date of the visit, the dollar amount spent, the services and products purchased, and any applied discounts.

The loyalty program does not have a membership fee. To become a member, the customer needs to provide an email address. Once in the program, members receive a $5-off reward coupon via email for every $100 they spend (across visits) on hair services and hair-care products. Given that an average transaction value is around $21, customers typically visit five times to earn a coupon. To redeem the coupons, members need to bring the coupon to the store or show the coupon email on a cellphone.[5]

In our data set, the price of a haircut increased by $1 one month after the loyalty program was launched, which is approximately a 5% increase. To normalize the analysis before and after the price increase, we add this price increase to the amounts spent before the price increase in order to avoid erroneously inferring that the loyalty program increases spending unless this increased spending is due to increased demand for services (or products). Because the price increase would potentially have a demand-reducing effect on customer visits, we view our findings as providing a lower bound on the program's impact on visit frequency and attrition. Alternatively, one can note that because the introduction of a loyalty program generally pairs a new program with the introduction of the discount from that program, assessing whether the success of a loyalty program is due to the presence of the program or the corresponding price discount is hard. In our case, any effect we find must come from the presence of the program itself because the total price to the customers in the post-program period is always at least as high as it is in the pre-program period.

Our empirical analysis leverages a unique group of loyalty-program members, whom we label *automatic members*. These customers provided their email addresses and agreed to receive marketing messages before the loyalty-program launch (and without awareness that a loyalty program would be introduced). When the loyalty program was introduced, the firm signed up these customers automatically for the program; therefore, the timing of joining the program is exogenous for these customers. In many settings, customers can decide when to join a loyalty program. In those cases, researchers should be concerned about the possibility of a dynamic selection bias: customers may select the timing of when to join the program in response to a shift in their demand for haircut services. Because of this self-selection, the behavior of a particular individual before and after he joins the loyalty program may reflect higher use of services, even while the program does not cause the demand increase. Using automatic members as the treatment group minimizes these dynamic selection bias concerns. Of course, other types of selection biases can apply to our analysis, and we discuss how we handle them in Section 2.2.

One clearly observable pre-program behavior in our quarterly cohort is the number of visits before the launch of the program. We leverage this observable source of heterogeneity to divide customers into frequency segments. Within each segment, we use automatic members as the treated group and non-members as the control group. Of course, it is possible that our matching approach does not control for differences between automatic and non-members in terms of the level of their concern for privacy, openness to receive marketing messages, or attitudes toward discounts, which could impact their response to the loyalty program. However, no information is available in our data set that can account for such differences, and we note this point as a caveat.

The customer cohort consists of 5,544 customers. Panel A of Table 1 shows the composition of the customer cohort. Approximately 38% of customers are members of the loyalty program, most of whom are automatic members who were enrolled by the firm at the program launch. About 2% of customers (or 5% of members) redeemed a reward coupon during our data period. Only these redeeming members entail costs to the company from running the loyalty program.[6]

Panel B of Table 1 shows the mean and standard deviation of the key variables of interest: net spending per visit (which takes into account any applied discounts), gross spending per visit (which reflects spending before accounting for applied discounts), days between visits, and the six-month hiatus rate, which we define for the purposes of the descriptive

Table 1. Summary Statistics Customer Cohort Composition

Panel A: Customer cohort composition

Types of consumers	Number of customers	Share of customer base	Number of transactions, mean (standard deviation)	Share of all transactions
All customers	5,544	100%	6.4 (7.5)	100%
Non-members	3,413	62%	4.9 (6.0)	47%
Members	2,131	38%	8.8 (8.8)	53%
Automatic members	1,769	32%	7.4 (8.3)	37%
Redeeming members	116	2%	23.6 (9.2)	8%

Panel B: Summary statistics of key variables of interest

Types of consumers	Net spending per visit ($), mean (standard deviation)	Gross spending per visit ($), mean (standard deviation)	Days between visits, mean (standard deviation)	Six-month hiatus rate, mean
All customers	20.63 (6.41)	20.66 (6.40)	50.14 (31.1)	17.83%
Non-members	20.18 (6.11)	20.19 (6.11)	54.44 (30.1)	24.35%
Members	21.03 (6.63)	21.09 (6.62)	46.83 (29.0)	12.02%
Automatic members	21.11 (6.38)	21.16 (6.37)	47.21 (29.8)	15.09%
Redeeming members	21.82 (7.50)	22.21 (7.41)	34.70 (18.6)	1.32%

analysis as an event where a customer does not visit for more than 182 days after a purchase occasion. The threshold of six months is somewhat arbitrary, but it is twice the length that the hair salon itself uses to decide that a customer has attrited, and very few men go more than six months between haircuts. Furthermore, although we use this length of hiatus for our descriptive analysis, we do not impose such a limit in our model-based analysis. Compared with non-members, automatic members spend more, visit more often, and are less likely to have a hiatus from visits for six months or longer during our observation window. We note, of course, that these summary statistics span both pre- and post-program introduction periods and therefore may reflect both the treatment effect of the loyalty program and the pre-program differences between automatic members and non-members.

2.2. Descriptive Analysis

In this section, we present some descriptive analysis about how the loyalty program affects different aspects of customer behavior: frequency of visits, spending per visit, and six-month hiatus rate. The purpose of this analysis is to demonstrate that the patterns of behavioral changes that we document in this paper are present in the underlying data and are not a result of the assumed model. Our modified HMM framework, discussed in the next section, focuses on changes in frequency and attrition and not on changes in spending because we find that the program has a negligible impact on spending in our empirical setting.

To quantify the change in behavior for automatic members after the program's introduction, it is important to find a proper control group. As is typical with observational data, the treatment (program membership) is not randomly assigned. Although they are enrolled into the loyalty program by the firm, automatic members nevertheless behave differently from non-members even before program introduction. For the descriptive analysis, we use three separate strategies to account for the selection effects. We describe those strategies first and then present the results of the analysis.

The first strategy is to match automatic and non-members by their number of salon visits before the program is introduced, as previously mentioned. We divide the customers into seven matched segments—automatic members and non-members who have one, two, three, four, five, six, and seven or more visits before the program introduction. We also use this approach to account for observed customer heterogeneity in our model (Section 3). The number of customers in each segment is shown in Table 2. We do not include the 321 customers who made seven or more trips to the firm before program introduction because these customers vary widely in the number of visits they made pre-program, and we have too few customers

Table 2. Number of Automatic Members and Non-members by Segment

Pre-program visit frequency	Automatic members	Non-members
1	685	1763
2	300	605
3	237	355
4	173	267
5	126	163
6	80	107
7+	168	153
Total	1,769	3,413

with any specific number of visits to conduct our full model analysis and get reliable estimates. The descriptive results are very similar if we include these customers, but for the purposes of comparison, we wish to be consistent and include the same customers in the descriptive and model-based analyses.

We then run the following regressions:

$$Y_{ik} = \beta_{ap} \cdot M_i \cdot AP_{ik} + \beta_m \cdot M_i + \beta_W \cdot W_k + s_i + s_i \cdot AP_{ik} + \epsilon_{ik}, \quad (1)$$

where Y_{ik} is the variable of interest for customer i on visit k, M_i is an indicator that equals one if customer i is an automatic member and zero if a non-member, AP_{ik} is an indicator that equals one if customer i's visit k occurs after program introduction, W_k is a vector of year-week dummies that represent a (nonparametric) common time trend (and seasonality) between automatic and non-members, s_i and $s_i \cdot AP_{ik}$ are the segment fixed effects (based on the number of pre-program visits), which can be different before and after program introduction, and ϵ_{ik} is an error term. The main parameter of interest β_{ap} is a difference-in-differences type of measure that is identified by how much Y changes for automatic members after program introduction compared with that for non-members.

The second strategy is to use customer fixed effects. Specifically, we run the following regression:

$$Y_{ik} = \beta_{ap} \cdot M_i \cdot AP_{ik} + \beta_W \cdot W_k + \alpha_i + \epsilon_{ik}, \quad (2)$$

where α_i is a customer fixed effect, and the other variables are as previously defined. In this specification, the individual fixed effects capture customer heterogeneity. The main parameter of interest β_{ap} is identified from the difference in within-person behavior change after program introduction among members compared with non-members.

The third strategy is to use propensity score matching. The goal is to find non-members who closely resemble the behavior of automatic members before program introduction to serve as the control group. We use customer data before the start of the program to run a logistic regression linking the customer's choice to give an email address (and thus become an automatic member) to relevant covariates. The covariates we use include the average spending, probability of product purchase, number of visits, and average number of visits per month before the program. For each member, we find the non-member with the closest propensity score as his control counterpart.[7] With the automatic members and matched non-members, we run the following regression:

$$Y_{ik} = \beta_{ap} \cdot M_i \cdot AP_{ik} + \beta_m \cdot M_i + \beta_W \cdot W_k + \epsilon_{ik}, \quad (3)$$

where M_i is an indicator variable that is set to one for automatic members and zero if customer i is a matched non-member. With propensity score matching, β_{ap} is identified from the difference in behavior change among members compared with matched non-members after program introduction.

2.2.1. Spending per Visit.

To analyze the impact of the program on spending, we compare both net and gross spending for automatic members before and after introduction of the loyalty program, using the spending per visit of non-members over time as a nonparametric control for time trends. We use each of the three strategies described previously to obtain a difference-in-differences measure of the program impact on spending for members.

Table 3 shows the estimation results. Columns (1)–(3) show the impact of the program on net spending per visit, which is the total amount the customer spends

Table 3. Spending per Visit

	Dependent variable: *Net Spending per Visit*			Dependent variable: *Gross Spending per Visit*		
Independent variables	Segment fixed effects (1)	Individual fixed effects (2)	Propensity score matching (3)	Segment fixed effects (4)	Individual fixed effects (5)	Propensity score matching (6)
After program start × automatic members (β_{ap})	−0.0494 (0.1144)	−0.1234 (0.1705)	0.2277 (0.1973)	0.0039 (0.1087)	−0.0805 (0.1705)	0.2865 (0.1972)
Automatic members	0.9887*** (0.2110)		0.1334 (0.1456)	0.9887*** (0.2109)		0.1335 (0.1455)
Number of visits before program dummies	Yes	No	No	Yes	No	No
Number of visits after program dummies	Yes	No	No	Yes	No	No
Individual dummies	No	Yes	No	No	Yes	No
Year + week dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22,231	22,231	16,174	22,231	22,231	16,174
R^2	0.0193	0.6209	0.0173	0.0202	0.6217	0.0176

$^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

after accounting for any rewards coupon he may have. Column (1) shows the results from Equation (1) with segment fixed effects, column (2) uses the individual fixed effects from Equation (2), and column (3) uses propensity score matching from Equation (3). The program's impact on net spending is neither statistically nor economically significant. Columns (4)–(6) show the same analysis for gross spending per visit, which is the amount the customer spends before accounting for any rewards coupon he may have. The coefficient β_{ap} remains small and insignificant, indicating that the loyalty program has a negligible effect on spending per visit. This insignificant impact on spending is likely a function of the industry we study, men's haircuts, where opportunities for the firm to increase the dollar value of a transaction are fairly limited.

2.2.2. Visit Frequency. We next assess the impact of the program on visit frequency, which we represent using the number of days between visits. A program effect that reduces the count of days between visits therefore implies a higher frequency of visits. Figure 1 shows a histogram of days between visits for customers who return within a year. Customers are most likely to revisit between 3 and 10 weeks after a previous visit. Because the data are right censored, we run our analysis on the number of days between visits conditional on customers returning within 182 days (which is 26 weeks, or approximately half a year). We use this limit only for the descriptive analysis, and it is neither necessary nor used for the HMM analysis in Section 3.

We conduct a regression analysis of visit frequency using the same three empirical specifications (Equations (1)–(3)), with Y_{ik} representing the number of days between the kth and $(k+1)$th visits for member i. We report the results in columns (1)–(3) of Table 4. The results from the three specifications for the frequency of visits (columns (1)–(3)) are very close. The program reduces the days between visits by approximately 2.3–2.9 days, which reflects an approximately 5%–6% increase in the frequency of visits.

2.2.3. Customer Hiatus. The final component we consider is how the loyalty program influences long-term customer hiatus, which we use as a proxy for customer attrition only for the purposes of descriptive analyses. We select a time duration of inactivity, 182 days (six months), after the last observed visit to denote the customer having a hiatus. The hiatus only partially accounts for unobserved customer attrition. Some of the increase in frequency of purchases may reflect fewer skipped visits for the focal firm (e.g., visits to a competitor). In our main HMM-based analysis in Sections 3 and 4, we model customer attrition directly and do not limit its definition to a period of inactivity after the last observed customer visit.

We cannot use individual fixed effects for the customer hiatus analysis because a six-month hiatus is not a repeated event for most customers. We therefore use the matched segments with the same number of visits and the propensity score matching approach to control for customer heterogeneity. We run a logistic regression model to estimate the program's effect on the probability of hiatus. The logistic regression analysis follows an analogous form as in Equations (1) and (3). In this case, Y_{ik} equals one if customer i does not revisit within 182 days after visit k and zero otherwise.

The results are shown in columns (4) and (5) of Table 4. After controlling for the time trend by the

Figure 1. Histogram of Number of Days Between Visits

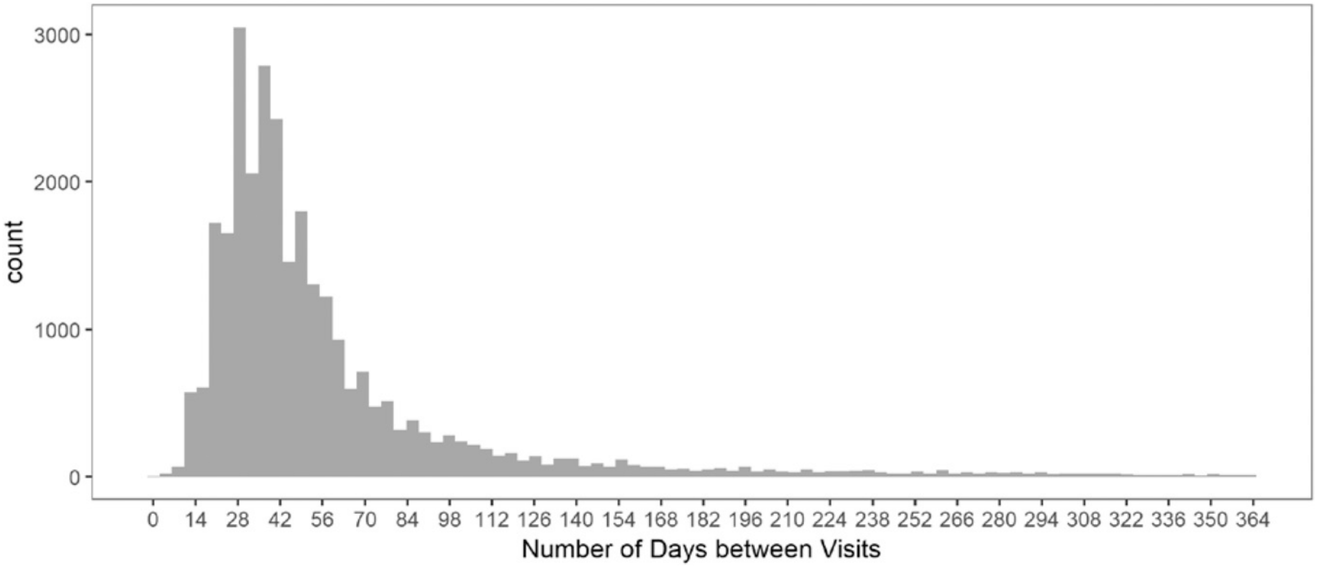


Table 4. Program Effect on Frequency (Days Between Visits) and Six-Month Hiatus

	Dependent variable: *Days Between Visits* (ordinary least squares)			Dependent variable: *Six-Month Hiatus* (logistic regression)	
Independent variables	Segment fixed effects (1)	Individual fixed effects (2)	Propensity score matching (3)	Segment fixed effects (4)	Propensity score matching (5)
After program start × automatic members	−2.9234*** (0.7366)	−2.3306* (1.4021)	−2.6092** (1.2323)	−0.3244*** (0.0968)	−0.2714*** (0.0879)
Automatic members	−2.0839*** (0.3137)		−1.8541* (0.9947)	−0.0123 (0.0735)	−0.0345 (0.0522)
Number of visits before program dummies	Yes	No	No	Yes	No
Number of visits after program dummies	Yes	No	No	Yes	No
Individual dummies	No	Yes	No	No	No
Year + week dummies	Yes	Yes	Yes	Yes	Yes
Observations	15,734	15,734	12,059	19,073	13,834
R^2	0.1866	0.4776	0.0649	0.4490	0.0719

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

year-week fixed effects, we see that automatic members do indeed have a lower probability of hiatus after program introduction using both approaches. To turn the parameter estimates into hiatus rates using the logistic function, we first take the sample average of all the other variables in the regression multiplied by their corresponding parameter estimates to get a baseline coefficient. For the specification with segment fixed effects, for example, the resulting baseline coefficient is −1.540, which comes from the weighted average of the year-week dummies and segment fixed effects as well as the dummy for automatic members. The baseline hiatus rate is then $\frac{\exp(-1.540)}{1+\exp(-1.540)} = 17.7\%$. For automatic members after program introduction, the hiatus rate is estimated to be $\frac{\exp(-1.540-0.168)}{1+\exp(-1.540-0.168)} = 15.3\%$. Thus, the hiatus rate reduces from 17.7% to 15.3% for automatic members after program introduction, which corresponds to a 13% reduction in relative terms. We apply the same procedure to the specification with propensity score matching. The hiatus rate reduces from 22.3% to 17.9% for automatic members after program introduction, or 19.5% in relative terms.

To summarize, the descriptive results demonstrate that the loyalty program has only a negligible effect on spending per visit and leads to about a 5%–6% increase in the frequency of visits. This finding is consistent with prior literature (Sharp and Sharp 1997, Lewis 2004, Leenheer et al. 2007). However, the loyalty program has a large impact on reducing customer hiatus (used as a proxy for attrition), which has been missed in the prior literature on non-tiered loyalty programs.

3. Model Development

We begin by discussing the need for a model to estimate the effects of a loyalty program on customer value in Section 3.1. In Section 3.2, we lay out the model framework. In Section 3.3, we discuss identification. In Section 3.4, we describe how the model is used to estimate the program effects.

3.1. Model Motivation

In order to estimate program effects on customer value, we seek to decompose the program's impact on the frequency of visits and attrition. We do not focus on changes in spending in our HMM setting given the negligible effects shown in Section 2.

Consider the example of a customer who had been visiting a salon approximately every four to six weeks but has been observed to have taken a long break from visiting. When projecting the stream of revenue one can expect from this customer, it would be prudent to consider the probability of this customer defecting, even if temporarily, to a competitor or other outside option that fulfills his need for a haircut, which means that our focal firm has zero probability of being visited, at least for some period of time. This defection constitutes (temporary) attrition, and failing to account for such defection can lead to a misestimation of each customer's value over a future time horizon.

A standard hazard model is not able to tease apart the temporary attrition and frequency effects of the program. In fact, a standard hazard model studies intervisit duration with the implicit assumption that there is no attrition. In other words, a standard model is likely to be misspecified in capturing the dynamic transitions that can lead a customer to become dormant (in which case attrition occurs) or resume patronage after a previous defection. These dynamics have implications for projecting customer revenue streams and ultimately for assessing the impact of a loyalty program.

The firm does not directly observe regime switches by the customer and can only infer the attrition probabilistically. We therefore propose a model in which we allow for latent states that capture being an active customer of the focal firm or having switched to a dormant state of inactivity with respect to the focal firm, where both states can be transient. An HMM (Netzer et al. 2008, Fader et al. 2010, Schweidel et al. 2011, Schwartz et al. 2014) is well suited for this purpose. A two-state HMM allows a customer to probabilistically switch between two regimes: being an active customer with some positive probability of visiting a salon and being a dormant customer who has no chance of visiting the salon. In our setting, attrition is defined as the customer moving from the active state to the dormant state. Two sets of parameters govern such an HMM: the state transition matrix (i.e., probabilities of toggling between the active and dormant states) and the probability of visiting while in the active state.

As the histogram of time between visits in Figure 1 demonstrates, the assumption that the chance of visiting a salon is independent of the time since the previous visit is not supported by the data. It is therefore important to allow the visit probability to change with the time spent in the active state since the previous visit (a property also known as *duration dependence*). This can be achieved by defining a visit hazard model for the active state. The transition probability matrix and the hazard function while in the active state collectively define the parameter space of this model, which we refer to as a *duration-dependent HMM* (DD-HMM). In Section 3.2, we define the modeling framework for the DD-HMM.

3.2. Model Definition

In our DD-HMM model, we have observable customer outcomes and latent customer states. The observable outcomes are visits denoted by an indicator variable Y_{it} that equals one if customer i visited the salon at time t and zero otherwise. Hence, each customer's outcomes can be represented as a vector of binary indicators $\boldsymbol{Y_i}$.

Our model has two states: active (A) and dormant (D). The state for a given customer i in time t is $s_{it} \in \{A, D\}$. We use variable d_{it} to denote the cumulative duration since entering the active state after the previous visit. In the active state, the probability that a customer visits the salon is positive and duration dependent: $p(Y_{it} = 1 | d_{it}, s_{it} = A) > 0$. By contrast, the customer will not visit the salon in the dormant state, that is, $p(Y_{it} = 1 | s_{it} = D) = 0$. For notational convenience, we define the discrete, duration-dependent hazard of visiting the salon conditional on being in the active state at time t as

$$h_{it}(d_{it}) = p(Y_{it} = 1 | d_{it}, s_{it} = A),$$

where the duration variable d_{it} is the number of days from either the last visit or the time in which the customer moved from the dormant to the active space (whichever is more recent). This variable is defined only for a customer in the active state, that is, $s_{it} = A$, as follows:

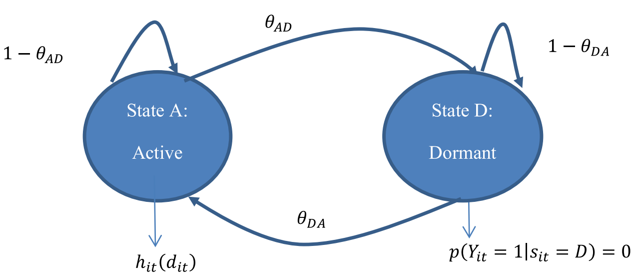
$$d_{it} = \begin{cases} 1, & \text{if } s_{i(t-1)} = D \text{ or } Y_{i(t-1)} = 1, \\ d_{i(t-1)} + 1, & \text{otherwise.} \end{cases}$$

The customer also has some probability of transitioning between states A and D. The probability of transitioning from state A to state D is denoted by θ_{AD}. A customer who is dormant can also return to the active state: the probability of transitioning from state D to state A is θ_{DA}. Customers start in state A at their first visit because their initial visit must occur in the active state such that $p(s_{i0} = A) = 1$. The framework is illustrated graphically in Figure 2. We also constrain the movement from the active to dormant state to only occur immediately after a purchase for reasons that are explained in Section 3.3.

The estimation strategy for the DD-HMM is to write the joint likelihood of the data and hidden states ($f(\boldsymbol{Y_i}, \boldsymbol{s_i}, \boldsymbol{d_i})$) and then marginalize this distribution to integrate out the hidden state sequence. As with a standard HMM, the DD-HMM has three components: the initial state distribution $p(s_{i0} = A)$, the state transition matrix defined by θ_{AD} and θ_{DA}, and the state-dependent choice probability $P(Y_{it} = 1 | d_{it}) = h_{it}(d_{it})$, where Y_{it} equals one if customer i chooses to visit at time t. As noted earlier, the initial state distribution is $p(s_{i0} = A) = 1$ because we assume that the customer enters the panel data with his first transaction, which can only happen in the active state A. The initial visit does not feature in the likelihood function.

We define state sequence $\boldsymbol{s_i^t} \equiv \{s_{i1}, \ldots s_{it}\}$ as a vector of states for customer i from time 1 to t. The joint likelihood of the data $\{Y_{it}\}$ and $\boldsymbol{s_i^T}$ for an individual customer i is

Figure 2. (Color online) Graphical Representation of the DD-HMM



$$
\begin{aligned}
&L\left(\theta_{AD}, \theta_{DA}, h_{it}(d_{it}) \mid \{Y_{it}\}, \boldsymbol{s}_i^T\right)\\
&\quad = \prod_{t=1}^{T} h_{it}\left(d_{it}\left(\boldsymbol{s}_i^t\right)\right)^{I(Y_{it}=1, s_{it}=A)} \cdot \left[1 - h_{it}\left(d_{it}\left(\boldsymbol{s}_i^t\right)\right)\right]^{I(Y_{it}=0, s_{it}=A)}\\
&\qquad \cdot (\theta_{AD})^{I(s_{i,t-1}=A, s_{it}=D, Y_{i,t-1}=1)} \cdot (\theta_{DA})^{I(s_{i,t-1}=D, s_{it}=A)}\\
&\qquad \cdot (1-\theta_{DA})^{I(s_{i,t-1}=D, s_{it}=D)} \cdot (1-\theta_{AD})^{I(s_{i,t-1}=A, s_{it}=A, Y_{i,t-1}=1)}.
\end{aligned} \tag{4}
$$

We note that the duration variables d_{it} are deterministic given a state sequence per its law of motion we outlined earlier. The marginal likelihood is obtained by summing over the space of possible state sequences $\boldsymbol{s}_i^T$, and the likelihood over all customers is formed by taking the product of the individual customer likelihoods. We maximize the log likelihood

$$
\begin{aligned}
&\ln(L(\theta_{AD}, \theta_{DA}, h_{it}(d_{it}) \mid \{Y_{it}\}))\\
&\quad = \sum_{q \in \boldsymbol{s}_i^T} \ln(L(\theta_{AD}, \theta_{DA}, h_{it}(d_{it}) \mid \{Y_{it}\}, q)).
\end{aligned} \tag{5}
$$

We next discuss identification of the model defined in Equation (5) and present exclusion restrictions required for identification.

3.3. Model Identification

In this section, we discuss why the general DD-HMM model defined in Section 3.2 is not estimable without certain exclusion restrictions. We then propose appropriate restrictions that allow for model identification.

A two-state HMM is identified if the state-dependent choice probability is independent of time spent in a given state (which is not the case for the DD-HMM, which allows for duration dependence). For example, suppose that the hazard function h_{it} is formulated as a function of exogenous covariates. The HMM is then identified because variations in visit probability not explained by these covariates are rationalized by (hidden) state switching.

However, in our model, a hazard function in the active state that changes with the time spent in that state leads to an identification challenge because the state sequence in an HMM is unobserved and has to be inferred. Allowing for an unrestricted hazard function and unrestricted state transitions leads to multiple explanations that can rationalize the same data. For example, if the hazard function could increase and then decrease after a certain number of weeks since the previous visit in an unconstrained manner, then any pattern of active versus dormant states could alternatively be explained through just a very flexible hazard function. Furthermore, the combination of unrestricted state switching and an unrestricted hazard could produce many sets of parameters that could each rationalize the same data.

This identification challenge can be resolved in two ways. First, we could impose the exclusion restriction that the hazard rate is always increasing (but otherwise let the hazard be nonparametric). Such an assumption would be justified in our industry because hair always grows longer over time, increasing the need for a haircut as time since the last haircut increases. Such an assumption might apply to most necessary goods. For example, the longer a customer has gone since he or she last went grocery shopping, the more he or she is likely to need to buy food. The same could be said about home-improvement products or appliances. The products where such an assumption would not be reasonable include leisure activities and many discretionary goods where the lack of a purchase does not lead to a future need. For example, the fact that a person does not have ice cream today does not necessarily increase the demand for ice cream tomorrow. A nonparametric hazard function, however, can overfit the data even if it is restricted to be increasing. A nonparametric hazard function can have as many parameters as the maximum unique intervisit duration in the data set. The number of data points for intervisit durations that are very short or long tend to be sparse, so the potential for overfitting is high.

The second approach is to use a parametric family of hazard functions, which leverages data across all intervisit durations to identify the parameters that best fit the data. Because one of our goals in conducting this model-based analysis is to construct forward-looking customer value simulations, we choose the approach of using a parametric hazard function for its greater parsimony. More specifically, we use a parsimonious parametric form called a *discrete Weibull model* (Nakagawa and Osaki 1975, Fader et al. 2018) for the hazard function in the active state that allows for sufficient flexibility in capturing the empirical patterns that we observe in the data:

$$
h_{it}(d_{it} \mid \alpha, c, p) = p \cdot \left(1 - (1-\alpha)^{d_{it}^c - (d_{it}-1)^c}\right).
$$

The discrete Weibull functional form offers a flexible set of discrete hazard functions whose shape depends on the parameter c. When $c > 1$, the hazard is increasing, when $c < 1$, the hazard is decreasing, and when $c = 1$ the hazard is flat. The hazard at time period 1 is $\alpha \cdot p$, and the steady-state hazard that the function asymptotes toward is p. With just three parameters (α, p, and c), the model can capture a variety of possible monotone hazard functions. Note that we no longer need to restrict the hazard to be increasing (i.e., by imposing $c > 1$) because this can be identified from the data (although we anticipate the increasing hazard given the previous logic). We ultimately use this parametric hazard function because it avoids the overfitting concerns of the nonparametric approach (which is particularly important for forward predictions) and fits the data well. In our empirical analysis, we find that $c > 1$, so the hazard rate is also monotonically increasing in our case.

We need one further restriction on the allowable state sequences between a pair of visits to the hair salon. For example, let us consider a customer who visited the salon four weeks after his previous visit. In week 4, the customer had to be in the active state because a visit was observed. During weeks 1–3, the customer could have exhibited the following $2^3 = 8$ possible state sequences (in which A refers to the active HMM state and D to the dormant HMM state of attrition), as shown in Table 5.

It is important to note that allowing for all these possible state sequences (i.e., unrestricted state switching) leads to identification issues because the data do not reveal whether a customer would have switched between states A and D repeatedly during any given pair of visits. We follow the prescription of Fader et al. (2005) to resolve this issue by restricting state transitions from the active to the dormant state to be allowed only in the first period after a visit. In other words, if the customer remains in the active state after the most recent visit, he is restricted to stay in the active state until the next visit. If the customer does shift to the dormant state in the first period after a visit, he can remain dormant or switch to active at any time period thereafter. Relating to Table 5, this means that the first four state sequences (*AAAA*, *DAAA*, *DDAA*, and *DDDA*) would be permitted under this restriction, whereas the last four would not (*ADAA*, *AADA*, *ADDA*, and *DADA*). Although we use the example of an intervisit duration of four periods, the identification issue is exacerbated when the duration between visits increases.

Operationalizing the transition from the active to the dormant state to only occur right after a visit (with probability θ_{AD}) is done by setting this transition to zero probability when the previous time period did not involve a visit. In other words, if a customer remains in the active state right after a visit, he stays there until the next visit, at which point he gets another opportunity to transition to the dormant state (similar to Fader et al. 2005). In the dormant state, customers can return to the active state with probability θ_{DA} in each time period.

In practical terms, the preceding state transition restriction is consistent with stylized customer behavior. We allow a customer to decide after a visit whether he wants to attrite from our focal firm and remain in that dormant state for some length of time before possibly returning to an active state of consideration.

With these restrictions, the DD-HMM is identified. The intuition of what identifies the state transition from the standard hazard model is this: the data on interpurchase time, as shown in Figure 1, have many shorter interpurchase times, as we would expect given the periodicity of men's haircuts. However, we also observe some very long interpurchase times. These observed times would be highly improbable under a monotonic hazard function.[8] These deviations would probabilistically be attributed as more likely to have come from attrition than from the hazard, with the longer the gap the higher the probability of that intervisit time being attributed to a defection to the dormant stage. The threshold at which a customer is inferred as having attrited is probabilistic and different for each segment.

Estimation of the DD-HMM model requires changes to the forward algorithm used in HMM estimation because the entire history of state transitions and outcomes between consecutive visits will matter in summing over the full set of paths that lead from the active state in which one visit occurred and the active state in which the next visit occurred. We describe a

Table 5. Possible State Sequences for a Duration of Four Periods Between Visits

State sequence	Duration in active state	Intuition
AAAA	Four periods	Always active since last visit
DAAA	Three periods	Dormant for one period and returned to active
DDAA	Two periods	Dormant for two periods and returned to active
DDDA	One period	Dormant for three periods and returned to active
ADAA	Two periods	Active for one period, then dormant for one period, and then returned to restart being active for two periods
AADA	One period	Active for two periods, then dormant for one period, and then returned to restart being active for one period
ADDA	One period	Active for one period, then dormant for two periods, and then returned to restart being active for one period
DADA	One period	Toggling between dormant and active every time period, ending up in the active state

novel and computationally efficient approach to estimating the DD-HMM in the appendix.

3.4. Using DD-HMM to Estimate Treatment Effects

The DD-HMM provides an integrated approach to model customer attrition and visit frequency. We now describe how we use the model to estimate the effect of program membership on attrition, visit frequency, and overall customer value. To estimate the change in θ_{AD}, the probability of attrition (transition from the active state to the dormant state), and the change in visit frequency (captured by hazard parameters α, p, and c, which we denote as a vector β for notational ease) after loyalty-program introduction,[9] we allow θ_{AD} and β to differ before and after program introduction. Mathematically, we define LP_{it} as an indicator of whether the program has launched at time t. As discussed in the appendix, we set $\theta_{AD}(LP_{it}) = \theta_{AD,before} \cdot \mathrm{I}(LP_{it} = 0) + \theta_{AD,after} \cdot \mathrm{I}(LP_{it} = 1)$ and $\beta(LP_{it}) = \beta_{before} \cdot \mathrm{I}(LP_{it} = 0) + \beta_{after} \cdot \mathrm{I}(LP_{it} = 1)$.

We are then able to obtain a difference-in-differences estimate of the program's effect for each attrition and weekly hazard probability, which represents how much of the automatic members' attrition and hazard can be attributed to the program effect. In this approach, non-members serve as an important control for time trends that are unrelated to membership.[10] To account for customer heterogeneity, we further segment automatic and non-members by the observed number of pre-program visits the customer made (six segments corresponding to one to six visits made before the loyalty program's introduction).[11] We then estimate separate models for each segment of customers, as shown in Table 6. We run separate models for each segment,[12] as well as separate models based on whether the customers are automatic members or non-members. Thus, our difference-in-differences estimates are

$$
\begin{aligned}
DID(\theta_{AD,s}) &= (\theta_{AD,after,auto,s} - \theta_{AD,before,auto,s}) \\
&\quad - (\theta_{AD,after,non,s} - \theta_{AD,before,non,s}) \text{ and} \\
DID(h_{it,s}) &= (h_{it,after,auto,s} - h_{it,before,auto,s}) \\
&\quad - (h_{it,after,non,s} - h_{it,before,non,s}),
\end{aligned} \tag{6}
$$

Table 6. Estimated DD-HMM Models

Segment (customer type)	Automatic members	Non-members
One pre-program visit	DDHMM-1A	DDHMM-1N
Two pre-program visits	DDHMM-2A	DDHMM-2N
Three pre-program visits	DDHMM-3A	DDHMM-3N
Four pre-program visits	DDHMM-4A	DDHMM-4N
Five pre-program visits	DDHMM-5A	DDHMM-5N
Six pre-program visits	DDHMM-6A	DDHMM-6N

where *auto* denotes automatic members and *non* denotes non-members.

Each of the models in Table 6 has a total of nine parameters (two sets of discrete Weibull hazard parameters for before and after program introduction and $\theta_{AD,before}$, $\theta_{AD,after}$, and θ_{DA}). This approach gives us a flexible way to account for consumer heterogeneity. In other words, instead of allowing for a random effect on the HMM parameters to capture heterogeneity based on a distributional form (which considerably increases computational complexity), we allow the parameters to be completely flexible for each segment. In total, 108 parameters are used across the entire data set to capture the behaviors of the six segments and two groups (automatic and non-members).

Finally, we seek to use the estimates to calculate the impact of the loyalty program on profits. We do this by simulating the discounted overall customer value over a five-year horizon for automatic members after program introduction and compare that to what the overall customer value would have been without the program. To simulate what these automatic members would have done had the program not been introduced, we subtract the difference-in-differences estimates of the attrition and hazard parameters denoted in Equation (6) from the parameters for automatic members after program introduction. This represents the appropriate counterfactual with parameters that maintain the same difference between automatic and non-members in the before-program and after-program periods. Furthermore, we are able to separate the relative effect of attrition and frequency on customer value by simulating the customer value under two further conditions: one in which only attrition effects are present (i.e., frequency effects are turned off) and one in which only frequency effects are present (i.e., attrition effects are turned off). The lift from each of these conditions provides an indication of whether attrition or frequency effects contribute more toward changes in customer value.

4. Model Results

In this section, we present the results from the estimated model. In Section 4.1, we show that the DD-HMM parameters for automatic members and non-members are generally similar before program introduction. In Section 4.2, we present the estimated effects of the program on attrition, frequency, and customer value. In Section 4.3, we present some robustness checks. The full DD-HMM model estimates are presented in the appendix.

4.1 Pre-program DD-HMM Parameters for Automatic and Non-members

Customers are segmented by pre-program visit frequency for both automatic members and non-members. We show here that the DD-HMM parameters in the pre-launch phase are similar between matched segments of automatic members and non-members. Although we do not require automatic members and non-members to exhibit no differences before the loyalty program is enacted because we are using a difference-in-differences approach, which merely requires a common set of time effects, the finding that our parsimonious matching approach appears to account for most of the difference between automatic and non-members before program introduction is reassuring. Full estimates of the DD-HMM model appear in the appendix. The key variables from the pre-program period include $\theta_{AD,before,m}$ (the transition probability from active to dormant) and the hazard parameters $\{\alpha_{before,m}, p_{before,m}, c_{before,m}\}$ for automatic members ($m = auto$) and non-members ($m = non$).

Table 7 presents the differences in these probabilities between automatic and non-members. None of the attrition probability parameters are statistically different between matched automatic and non-member segments, which suggests that non-members do not have a higher attrition probability than automatic members before program introduction. Most of the hazard parameters are not statistically different except for segment 6 (which is comprised of consumers who visited the salon six times before the loyalty program was introduced). We show the mean hazard in the pre-program phase for automatic members versus non-members in Figure 3, which suggests that although the pre-program hazards are statistically different for segment 6, the magnitude of the difference is relatively small (comparing the distance between the two lines in Figure 3 at each duration). Therefore, our estimation strategy controls for most of the pre-program differences between automatic members and non-members. Our difference-in-differences approach will also account for the pre-program differences in the hazard for segment 6.

4.2. Program's Effect on Attrition, Frequency, and Overall Customer Value

In this section, we analyze the program's effects on customer value. First, we examine the program's effect directly on the attrition probabilities. We then calculate the overall effect of the program on customer value and break that effect down into how much comes from attrition and frequency, demonstrating that most of the effect comes from decreased attrition.

We begin by discussing the direct difference-in-differences estimates of the program effect on attrition and visit frequency. First, we examine how the program changed the attrition rate, which is measured as how much θ_{AD} is reduced for automatic members after program introduction. For this and all subsequent analyses, we take draws from a normal distribution using the parameter estimates and standard errors in order to account for parameter uncertainty. The estimated reduction in the attrition probability is given in Table 8. Overall, we find that the attrition rate decreases by 5.4 percentage points by taking a weighted average of the segment-specific effect sizes. We estimate that the overall attrition probability without the program would be approximately 23.3%. Thus, taking the ratio of these numbers enables us to find that the loyalty program leads to a 23% (=5.4/23.3) relative reduction in attrition probability (across the customer base). This result is higher than that in the descriptive analysis of Section 2, which found a relative reduction in the range of 13%–19.5% in attrition rate. The main reason for this difference lies in the more appropriate behavior-based definition of attrition used in our modeling approach compared with the threshold of 182 days of inactivity used to define attrition in Section 2. Specifically, our model picks up attrition that occurs when a customer skips one or two haircuts but then returns to our focal hair salon within 182 days. This is especially relevant for previously frequent customers. Such an event would be attributed as a change in frequency in the regressions in Section 2 but is better described as (temporary) attrition.

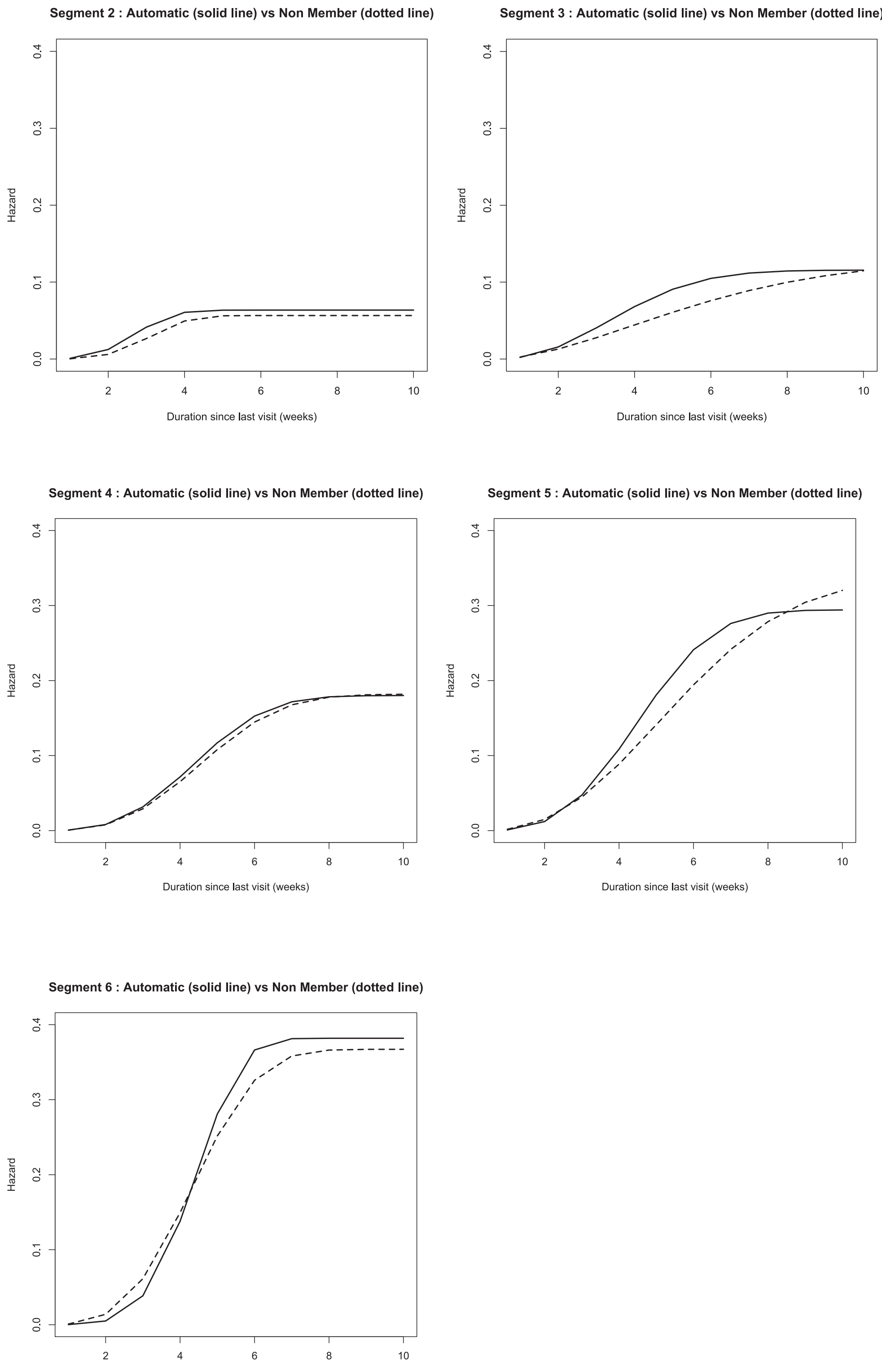
The overall attrition effect, however, masks heterogeneity of the result. The program has the largest effect of reducing attrition on the least (one visit prior to program) and most (six visits prior to program) frequent customers. We do not see a statistically significant effect for moderately frequent customers (with two to five pre-program visits). These results

Table 7. Difference in DD-HMM Attrition and Hazard Parameters Before Program Introduction

	Segment (pre-program visit frequency)					
Variable	1	2	3	4	5	6
$\theta_{AD,before,auto} - \theta_{AD,before,non}$	−0.014	0.048	0.033	−0.000	−0.009	−0.003
$\alpha_{before,auto} - \alpha_{before,non}$	n/a	0.007	−0.002	−0.000	−0.002	**−0.002**
$p_{before,auto} - p_{before,non}$	n/a	0.007	−0.014	−0.002	−0.042	0.015
$c_{before,auto} - c_{before,non}$	n/a	−0.337	0.524	0.088	0.623	**1.217**

Notes. Differences that are statistically significant ($p < 0.05$) are in boldface. n/a, not applicable.

Figure 3. Pre-program Hazard Function for Segments 2–6 Comparing Automatic Members (Solid Line) and Non-members (Dotted Lines)



Note. Segment 1 has a zero hazard because the segment 1 customers do not have a repeat transaction.

highlight the following. First, segment 1 is composed of customers who have made only one pre-program visit to the salon and have experienced a hiatus of at least six months. Of those customers who do return after program launch, we find that automatic members have an attrition rate that is 7.2 percentage points

Table 8. Difference-in-Differences Program Effect Estimate on Attrition (Probability of Transition from Active to Dormant State)

Segment	Program estimate	Standard error	p-Value
One pre-program visit	−0.072	0.029	**0.013**
Two pre-program visits	−0.062	0.039	0.114
Three pre-program visits	−0.051	0.031	0.101
Four pre-program visits	0.001	0.023	0.717
Five pre-program visits	−0.030	0.024	0.219
Six pre-program visits	−0.052	0.020	**0.011**

Note. Differences that are statistically significant ($p < 0.05$) are in boldface.

lower than the rate for non-members. This finding suggests that the loyalty program reduces the chance of entering a long hiatus even for the segment 1 customers who will take a long time to receive a reward coupon at their rate of visits. Second, segment 6, which is composed of highly frequent customers, experiences an absolute reduction in attrition of 5.2 percentage points. This finding suggests that the program is also effective at reducing periods of hiatus for these frequent customers, perhaps because these customers anticipate that they are likely to earn rewards.

We next synthesize how these elements add up in terms of the total value the firm gets from its customers. We compute the effect of the program on overall customer value over a five-year (or 260-week) horizon (similar to Fader et al., 2010), which we deem as a managerially relevant benchmark. We use the DD-HMM parameter estimates for each segment to forward simulate customer visits on a weekly basis and compute the discounted expected number of visits from the simulated data. We apply discounting at the weekly level with a nominal annual discount rate of 10%.[13] We then multiply the discounted expected number of visits by a $21 price per visit (which is the most common amount as the price of a basic hair service) to obtain the overall five-year discounted expected customer value.

We calculate the impact of the loyalty program using two sets of simulations for each segment. In the first set, we use the DD-HMM parameters for automatic members in the post–program introduction period, which includes the full effect of the program on attrition and frequency. In the second set, we use the same parameters as in the first set less the corresponding difference-in-differences estimates for attrition probability and the hazard function parameters, which represent a control condition of what the automatic members' parameters would have been in the absence of the estimated program effects.

We show the results in Table 9. Column (1) reports the dollar value the firm would earn per customer in each segment if the firm did not have the loyalty program. Column (2) reports how much more the firm is able to extract from consumers in each of the segments and in total because of the presence of the loyalty program. Finally, column (3) reports the percentage increase in customer value as a result of the loyalty program. The program increases customer value for all segments but in a U-shaped pattern in the point estimates as a function of pre-program visit frequency. However, the impact is statistically significant ($p < 0.05$) only for segments 1 and 6 in terms of revenue lift (in both dollar and percentage terms), and these lifts are generally not statistically significantly different from each other. These results, which are similar to the effect of the program on the attrition parameter (shown in Table 8), show that program membership, although providing a healthy return for highly frequent customers (i.e., segment 6) over a five-year horizon, is also effective at generating lift for the most infrequent customers.

Table 9. Program Effect and Lift Estimates on Discounted Expected Five-Year Customer Value

Segment	Customer value without program (standard error)	Change in customer value (standard error)	Lift in customer value (standard error)
1	$72.43 ($7.53)	$23.18** ($10.28)	32.9%** (15.5%)
2	$114.12 ($17.35)	$41.50** ($20.06)	38.7%* (22.0%)
3	$144.93 ($17.44)	$39.77 ($25.95)	29.2% (20.9%)
4	$251.67 ($38.93)	$3.36 ($42.90)	3.5% (18.0%)
5	$301.38 ($43.34)	$51.19 ($48.19)	19.0% (18.0%)
6	$315.69 ($44.38)	$119.72** ($50.24)	39.8%** (19.2%)
Overall	$140.52 ($7.79)	$33.95*** ($9.15)	29.5%*** (8.9%)

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

The highly frequent segment 6 has the most economic incentive to stay with the firm because these customers are more likely to earn rewards with repeated visits under the program and is likely a driver for the large effect found for this segment. However, such an incentive is unlikely to drive the behavior of segment 1 customers, who, by definition, had a hiatus of at least six months when the program was introduced. The attrition reduction for this segment suggests that psychological drivers could be at play for this segment. Based on this analysis, conferring program membership appears to have an effect even for those low-frequency customers who are less likely than high-frequency customers to earn a coupon for redemption.

The total lift in customer value is 29.5%, which represents a substantial gain in average revenue per customer over five years and is highly statistically significant. In Online Appendix A, we show a robustness check by estimating a DD-HMM that also allows for unobserved heterogeneity within each segment. We find an approximately 30% lift in overall customer value, which is primarily driven by segments 1 and 6, similar to Table 9.

To understand the relative impact of attrition versus frequency effects on overall customer value, we run the same comparison as earlier, except that we turn off the frequency effects by setting them to their implied levels if the program had not existed. Thus, we measure how much of an increase in customer value can be directly attributed to reduced attrition. The results are shown in Table 10. They demonstrate a pattern similar to the preceding one, which is that the largest effects from attrition reduction occur for the least and most frequent customers, although only the estimate from segment 4 is statistically significantly different from the estimates of segments 1 and 6. The aggregate lift of the program from reduced attrition is 23.6%, which represents about 80% of the 29.5% total lift of the program.

We also conduct a similar analysis about the impact of the program on frequency effects. In this case, we turn off the attrition effects of the program and then compare how average customer value changes when customers have the hazard rates of visits with the program compared with what their hazard rate would have been if the program had not been implemented. The results are shown in Table 11. We observe the aggregate lift due to frequency-only effects is 4.1% (compared with the total lift of 29.5%). However, these effects are all statistically insignificant, so all we can conclude is that, in general, the effect of the loyalty program on frequency is small, at least compared with the effect of the program on attrition. This may perhaps be due to the nature of the industry we study, men's haircuts, which may have an optimal interpurchase period for many customers.

In summary, we see that the program significantly increases customer value for the firm, obtaining a 29.5% lift. Most of this effect comes from attrition (about 80%), whereas a much smaller proportion comes from the increased frequency of visits. We also note that the total effect of the program on customer value, as measured in Table 9, is slightly larger than the combined effect of the program reported in Table 10 (attrition) and Table 11 (frequency). The reason is that a complementarity exists in attrition and frequency in increasing customer value: increased attrition becomes more valuable if customers come more frequently. In other words, the value of increased frequency is only beneficial if the customers actually stay.

4.3. Robustness Checks

We present three robustness checks to the analysis presented in the preceding section. First, we present a DD-HMM model that allows for within-segment unobserved heterogeneity. Second, we explore how customer value lift varies as a function of the time horizon used. Third, we allow for spend per visit to be drawn from segment-specific distributions.

In Online Appendix A, we present modeling modifications that allow for a latent mixture of DD-HMMs to be estimated for each segment and group (automatic and non-members). This model allows each customer in a segment to belong to one of two latent mixtures,

Table 10. Attrition-Only Estimates on Discounted Expected Five-Year Customer Value

Segment	Customer value without program (standard error)	Change in customer value (standard error)	Lift in customer value (standard error)
1	\$72.43 (\$7.53)	\$22.37** (\$10.02)	31.9%** (15.2%)
2	\$114.12 (\$17.35)	\$26.06 (\$16.45)	24.9% (16.9%)
3	\$144.93 (\$17.44)	\$29.18 (\$18.02)	21.0% (13.8%)
4	\$251.67 (\$38.93)	–\$11.35 (\$25.95)	–3.7% (9.7%)
5	\$301.38 (\$43.34)	\$33.64 (\$30.50)	11.8% (10.4%)
6	\$315.69 (\$44.38)	\$97.37** (\$39.92)	31.8%** (14.2%)
Overall	\$140.52 (\$7.79)	\$25.06*** (\$6.88)	23.6%*** (7.9%)

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Table 11. Frequency-Only Estimates on Discounted Expected Five-Year Customer Value

Segment	Customer value without program (standard error)	Change in customer value (standard error)	Lift in customer value (standard error)
1	\$72.43 (\$7.53)	\$0.26 (\$4.80)	0.0% (6.5%)
2	\$114.12 (\$17.35)	\$10.15 (\$8.19)	8.8% (6.8%)
3	\$144.93 (\$17.44)	\$7.95 (\$14.68)	6.0% (11.2%)
4	\$251.67 (\$38.93)	\$14.41 (\$31.96)	6.9% (13.5%)
5	\$301.38 (\$43.34)	\$14.31 (\$35.68)	5.8% (12.5%)
6	\$315.69 (\$44.38)	\$12.42 (\$29.63)	4.6% (9.7%)
Overall	\$140.52 (\$7.79)	\$6.49 (\$5.81)	4.1% (4.0%)

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

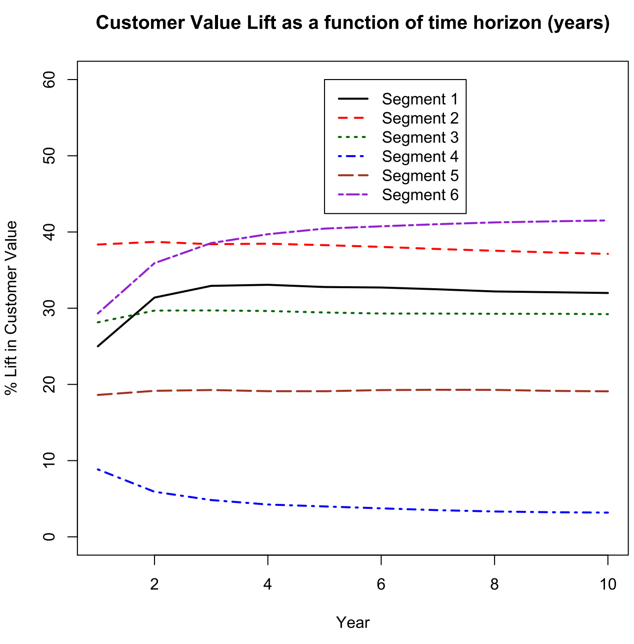
each having its own transition probabilities and hazard functions. This extension could be useful if, within count segments, there are customers of different subtypes who may have differing behavioral responses to the loyalty program. We show in Online Appendix A that the primary findings from Section 4.2 continue to hold when controlling for unobserved heterogeneity. That is, segments 1 and 6 continue to drive customer value improvement, and the overall customer base experiences about a 30% lift in value because of program introduction. This suggests that the count segments in Section 4.2 are relatively effective in capturing average program treatment effects.

Our second robustness check concerns the sensitivity of our results to analyzing the CLV lift using a five-year horizon on customer value. In Figure 4, we plot the percentage lift in customer value (a metric we report in Table 9 using a forward time horizon of five years) as a function of time horizon ranging from 1 to 10 years. Although an infinite-horizon CLV has conceptual value, most businesses have planning horizons that are much shorter than infinite time, and customer value eventually converges to steady state based on an HMM.[14] From Figure 4, we observe that the lift in customer value converges over the 10 years for each of the segments and that the lift for a 5-year horizon is very close to a 10-year one. That is, most of the dynamics in the model are shaken out within the first few years of forward simulation. Our results are therefore robust to increases in time horizon.

Finally, Online Appendix C presents the customer value lift that is obtained when we allow for spending amounts to be drawn from segment-specific distributions and find negligible changes from what we report in Table 9. Therefore, our results are robust to variation in spend amounts.

Overall, our measured impact of the loyalty program is much higher than the impact other papers studying these programs have found. One reason for this is that other papers have not been able to measure the impact of the program on attrition rates, which we show is a key driver of the loyalty program's value.

Figure 4. (Color online) Customer Value Percentage Lift as a Function of Time Horizon from 1 to 10 Years



5. Discussion and Conclusion

In this study, we estimate the effect of a non-tiered loyalty program on customer value over a five-year time horizon and decompose the drivers of this change in customer value into effects from attrition, visit frequency, and monetary spending. We show through careful descriptive analysis that the program has a negligible effect on the amount spent per visit. Both the descriptive analysis and a model-based approach demonstrate that the program has only a small effect on visit frequency and a large effect on attrition prevention. Although we believe that our model provides the best measures of frequency and attrition, the finding that the magnitudes of the descriptive and modeled results are similar is reassuring because it suggests that the effects come directly from the data and are not imposed by the model structure. Ultimately, we find that the overall customer value over a five-year horizon improves by 29.5%, primarily driven by the lower attrition rates as a result of the program.

We also find that the effects are largest for customers who are most and least involved with the company before the program is implemented. This result might be surprising given the conventional wisdom that loyalty programs have the most benefit to moderately involved customers. The logic of why moderately involved customers might respond the most is that they have room to grow with the company and can be motivated by the extra rewards they earn, whereas the customer who rarely comes in will have too low a probability of redemption to be motivated by the program, and the most frequent customers will earn the rewards regardless of their behavior on a particular visit. This logic is driven by an up-sell and frequency mentality, and indeed, we see that the moderate customers are the ones who increase their frequency the most (although the result is not statistically significant). However, we find that it is the extreme groups whose loyalty is most affected by the retention effects of the program, which dominate the economic impact of the program. Thus, the previous focus of the rewards literature on non-retention aspects of rewards programs may also affect the conventional wisdom of which customers to target.

Although the measured effect comes from the empirical context of a hair salon chain, we expect the insight that loyalty programs can create a large value through reducing the attrition rate to carry through to other settings. At the same time, the small effects on spending and frequency, although consistent with previous literature, can be specific to the hair salon industry, where the opportunities for up-selling are relatively limited and the demand in the category is tied to natural cycles of hair growth.

Multiple theories, including psychological reasons and economic incentives, could jointly contribute to the overall effect. On the one hand, the shifting behaviors around a reward, such as increasing visit frequency when getting close to using a reward (results are shown in Online Appendix B), could be consistent with economic incentives. On the other hand, only 2% of customers redeem a reward coupon. Furthermore, even low-frequency customers show a behavior change in terms of attrition when enrolled in the program. These results suggest that the benefit of the program can extend beyond economic factors and that psychological benefits also arise for members in the loyalty program. There are several psychological reasons why this program might increase loyalty: the program could provide an idea that the customer has chosen to make a relationship with the firm (consistency; e.g., Fishbach et al. 2011); the program could create the perception that the company is investing in its customers (De Wulf et al. 2001) or yield other forms of brand affinity (Thomson et al. 2005), and that could, in turn, generate goodwill for the company; or it could give a sense that by choosing to belong to the program, the customer can unlock special benefits, conferring a status to the customer (Lacey et al. 2007). We leave it to future research to tease apart and decompose the overall impact from different mechanisms.

Our study makes three significant contributions to the literature. First, we find a larger overall effect for a non-tiered loyalty program than has previously been found in the literature. This larger effect is attained by accounting for the impact on customer attrition—an aspect that extant literature on the effectiveness of loyalty programs has largely been silent on. Second, we leverage variation in our data that minimizes selection biases to the extent possible, which is aided significantly by having data from the customers both before and after the loyalty program is introduced and the availability of members who were signed up to the program by the firm (automatic members). Third, separating attrition and visit-frequency effects is non-trivial because attrition is unobserved in the non-contractual setting, which is the case in our empirical context. Hence, we implement a modeling approach that is an extension of an HMM that allows for a duration-dependent hazard function while the customer actively considers visiting the hair salon and a zero probability of a visit when the customer is in a dormant state. Our model therefore is able to separate out these effects to better understand how much the program affects attrition and frequency and to use simulations based on estimated parameters to compute the five-year lift in customer value. The suggested HMM can also be useful beyond the current application to measuring loyalty program effectiveness because marketing managers can use the inferred knowledge about customer attrition for development of marketing tactics.

Acknowledgments

The authors thank Eric Bradlow, Chakravarthi Narasimhan, and attendees at seminars at Emory, Johns Hopkins, and Harvard Universities, the National University of Singapore, Rochester University, Washington University in St. Louis, the New Economic School, the 11th Triennial Choice Symposium, Marketing Science (Rome), Summer Institute in Competitive Strategy, Frank M. Bass – UT Dallas Frontiers of Research in Marketing Science (UTD FORMS) Conference, and Marketing Dynamics Conference (Hong Kong) for helpful insights in shaping this paper.

Appendix. Full Set of DD-HMM Parameter Estimates

Model Estimation Algorithm

We estimate the DD-HMM described in Section 3 using maximum likelihood estimation (MLE). A standard HMM could be estimated using the forward algorithm, which integrates over the possible state sequences to arrive at an

Table A.1. DD-HMM Parameters for Automatic Members

	Segment (pre-program visit frequency)					
Variable	1	2	3	4	5	6
$\theta_{DA,auto}$	0.003	0.006	0.008	0.010	0.017	0.012
$\theta_{AD,before,auto}$	0.957	0.369	0.182	0.089	0.053	0.030
$\theta_{AD,after,auto}$	0.230	0.169	0.161	0.117	0.114	0.064
$\alpha_{before,auto}$	n/a	0.008	0.016	0.003	0.003	0.000
$p_{before,auto}$	n/a	0.064	0.116	0.180	0.291	0.384
$c_{before,auto}$	n/a	4.207	3.158	3.934	3.848	5.465
$\alpha_{after,auto}$	0.014	0.004	0.002	0.002	0.001	0.001
$p_{after,auto}$	0.132	0.143	0.198	0.269	0.419	0.421
$c_{after,auto}$	3.508	4.097	4.097	3.806	4.692	5.075

overall data likelihood for a given set of parameters. The forward algorithm works because of the first-order Markovian property of a standard HMM: that only the previous state affects the current state. However, as illustrated in Table 5, the DD-HMM also features duration since last visit d_{it} as a state variable, and this variable does not accumulate when in the dormant state. As a result, it becomes necessary to integrate over possible values of d_{it} in the likelihood function, which requires keeping track of more than just the previous state (active or dormant).

We therefore develop a modification of the forward algorithm that exploits the fact that each customer's data can be represented as a series of inter-visit durations (after the initial visit) and a survival period at the end that represents how long the customer has not yet visited. We therefore need to take into account all possible state sequences that can occur for a given inter-visit or survival duration. The number of such possible sequences increases rapidly as the inter-visit duration increases.

For computational efficiency, we pre-compute the log likelihood of various components of the likelihood function for different inter-visit and survival times so that these are not recomputed for each individual. What helps our model be computationally efficient is that parameters for a given customer segment are homogeneous (i.e., there are before- and after-program DD-HMM parameters that apply to all customers within a segment), whereas heterogeneity is flexibly captured by estimating separate models for each segment.

Table A.2. DD-HMM Parameters for Non-members

	Segment (pre-program visit frequency)					
Variable	1	2	3	4	5	6
$\theta_{DA,non}$	0.002	0.004	0.006	0.005	0.017	0.016
$\theta_{AD,before,non}$	0.969	0.324	0.148	0.088	0.059	0.034
$\theta_{AD,after,non}$	0.314	0.184	0.180	0.111	0.152	0.121
$\alpha_{before,non}$	n/a	0.003	0.019	0.003	0.005	0.002
$p_{before,non}$	n/a	0.056	0.128	0.182	0.337	0.363
$c_{before,non}$	n/a	4.684	2.634	3.815	3.321	4.158
$\alpha_{after,non}$	0.022	0.012	0.002	0.002	0.001	0.000
$p_{after,non}$	0.123	0.103	0.164	0.209	0.369	0.462
$c_{after,non}$	3.238	3.425	3.846	4.009	4.113	5.521

Note. n/a, not applicable.

Model Parameters for a Given Customer Segment

The parameters for a given customer segment (within either automatic or non-members) are the set of state transition and hazard parameters (as represented by a discrete Weibull model) before and after program introduction:

$$\theta_{AD}(LP_{it}) = \theta_{AD,before} \cdot I(LP_{it} = 0) + \theta_{AD,after} \cdot I(LP_{it} = 1) \text{ and}$$
$$\beta(LP_{it}) = \beta_{before} \cdot I(LP_{it} = 0) + \beta_{after} \cdot I(LP_{it} = 1).$$

The transition probability going from dormant to active state θ_{DA} is not a function of program introduction because there is not enough variation in the data to estimate these before and after programs. Hence, a single θ_{DA} is estimated across the entire time horizon.

For the hazard parameters, we have three parameters under the discrete Weibull model for the before-program hazard and another three parameters for the hazard after the program is launched.

Estimated Parameters

We use MLE to obtain parameter estimates for each DD-HMM. We present the full set of parameter estimates for the models for automatic members (Table A.1) and non-members (Table A.2). Segment 1 does not have a pre-program hazard defined because it has no repeat visits after an initial transaction.

We compute standard errors from the Hessian matrix estimated during MLE and simulate parameter draws from the distribution with the MLE parameters as the mean and the covariance matrix based on the Hessian. The discrete Weibull parameter draws are then transformed into hazard probability draws. These draws are used to compute difference-in-differences estimates iteration by iteration by taking the difference between the after and before automatic member parameters and the after and before non-member parameters. In order to compute what the automatic member's parameters would have been in the absence of a program effect, we subtract the difference-in-differences estimate from the automatic member's after parameters.

Endnotes

[1] Taylor and Neslin (2005) also find that a similar program that rewards shoppers with a turkey for meeting a monetary spend threshold increased sales by about 6% during an eight-week period and by 1.8% over the seven weeks after the promotional period. Using an estimate of costs in the range of 2%–4% based on the data in this paper, the benefit from the program would be in a range similar to that of the other papers listed here. The limited time nature of these programs may also affect the size of the measured impact.

[2] We use the term CLV to denote a five-year time horizon in our data setting, as explained in Section 4.2.

[3] Although non-members' decision not to join the loyalty program could be strategic, we show that matching automatic members and non-members on visit frequency before the program results in very similar pre-program behaviors.

[4] These customers do not have a visit during the six-month window before their observed first visit in our data set. Hence, we assume that these customers are newly acquired during their first observed visit.

[5] Some customers may fail to redeem available reward coupons (which expire in 90 days) because of forgetting or taking another discount that cannot be combined with the coupon.

[6] The program's costs are all fixed, except for the redemption costs.

[7] We pick 0.001 as the maximum distance for the probability between a matched member and non-member.

[8] If one chose to use parametric identification, the same would be true for deviations from the assumed form of the hazard rate.

[9] We note that the treatment is a regime change (launch of the loyalty program) rather than a temporary benefit that expires.

[10] Examples can include opening of competing stores or changes in hair style preferences (e.g., keeping one's hair longer or shorter).

[11] The number of customers with seven or more visits before program introduction was too low to reliably estimate treatment effects.

[12] By definition, segment 1 has no repeat visits in the pre-program phase. Therefore, its hazard function parameters are not identified, and the pre-program hazard is set to zero for both automatic members and non-members.

[13] Discounting is done at the weekly level because it is the unit of time in our model-based analysis. The weekly discount rate is therefore 10%/52 and is compounded on a weekly basis to better reflect when cash flows are incurred.

[14] This is because any Markov chain eventually converges to a unique stationary distribution across states.

References

Ascarza E, Hardie BGS (2013) A joint model of usage and churn in contractual settings. *Marketing Sci.* 32(4):570–590.

Berry J (2013) The 2013 COLLOQUY loyalty census. COLLOQUY industry report. Accessed December 15, 2015, http://www.colloquy.com/files/2013-COLLOQUY-Census-Talk-White-Paper.pdf.

Bolton RN, Kannan PK, Bramlett MD (2000) Implications of loyalty program membership and service experiences for customer retention and value. *J. Acad. Marketing Sci.* 28(1):95–108.

De Wulf K, Odekerken-Schröder G, Iacobucci D (2001) Investments in consumer relationships: A cross-country and cross-industry exploration. *J. Marketing* 65(4):33–50.

Drèze X, Nunes JC (2009) Feeling superior: The impact of loyalty program structure on consumers' perceptions of status. *J. Consumer Res.* 35(6):890–905.

Fader PS, Hardie BGS, Lee KL (2005) "Counting your customers" the easy way: An alternative to the Pareto/NBD model. *Marketing Sci.* 24(2):275–284.

Fader PS, Hardie BGS, Shang J (2010) Customer-base analysis in a discrete-time non-contractual setting. *Marketing Sci.* 29(6):1086–1108.

Fader PS, Hardie BGS, Liu Y, Davin J, Steenburgh T (2018) "How to project customer retention" revisited: The role of duration dependence. *J. Interactive Marketing* 43(1):1–16.

Fishbach A, Ratner RK, Zhang Y (2011) Inherently loyal or easily bored? Nonconscious activation of consistency versus variety-seeking behavior. *J. Consumer Psychol.* 21(1):38–48.

Hartmann WR, Viard BV (2008) Do frequency reward programs create switching costs? A dynamic structural analysis of demand in a reward program. *Quantitative Marketing Econom.* 6(2):109–137.

Henderson CM, Beck JT, Palmatier RW (2011) Review of the theoretical underpinnings of loyalty programs. *J. Consumer Psychol.* 21(3):256–276.

Kopalle PK, Sun Y, Neslin SA, Sun B, Swaminathan V (2012) The joint sales impact of frequency reward and customer tier components of loyalty programs. *Marketing Sci.* 31(2):216–235.

Lacey R, Suh J, Morgan RM (2007) Differential effects of preferential treatment levels on relational outcomes. *J. Service Res.* 9(3):241–256.

Lal R, Bell DE (2003) The impact of frequent shopper programs in grocery retailing. *Quantitative Marketing Econom.* 1(2):179–202.

Leenheer J, van Heerde HJ, Bijmoltc THA, Smidtsd A (2007) Do loyalty programs really enhance behavioral loyalty? An empirical analysis accounting for self-selecting members. *Internat. J. Res. Marketing* 24(1):31–47.

Lewis M (2004) The influence of loyalty programs and short-term promotions on customer retention. *J. Marketing Res.* 41(3):281–292.

Liu Y (2007) The long-term impact of loyalty programs on consumer purchase behavior and loyalty. *J. Marketing* 71(4):19–35.

Nakagawa T, Osaki S (1975) The discrete Weibull distribution. *IEEE Trans. Reliability* 24(5):300–301.

Netzer O, Lattin JM, Srinivasan V (2008) A hidden Markov model of customer relationship dynamics. *Marketing Sci.* 27(2):185–204.

Orhun AY, Guo T (2019), Reaching for gold: Frequent-flyer status incentives and moral hazard. Working paper, University of Michigan, Ann Arbor, MI.

Schmittlein DC, Morrison DG, Colombo R (1987) Counting your customers: Who are they and what will they do next? *Management Sci.* 33(1):1–24.

Schwartz EM, Bradlow ET, Fader PS (2014) Model selection using database characteristics: Developing a classification tree for longitudinal incidence data. *Marketing Sci.* 33(2):188–205.

Schweidel DA, Bradlow ET, Fader PS (2011) Portfolio dynamics for customers of a multiservice provider. *Management Sci.* 57(3):471–486.

Sharp B, Sharp A (1997) Loyalty programs and their impact on repeat-purchase loyalty patterns. *Internat. J. Res. Marketing* 14(5):473–486.

Stourm V, Bradlow ET, Fader PS (2015) Stockpiling points in linear loyalty programs. *J. Marketing Res.* 52(2):253–267.

Taylor GA, Neslin SA (2005) The current and future sales impact of a retail frequency reward program. *J. Retailing* 81(4):293–305.

Thomson M, MacInnis DJ, Whan Park C (2005) The ties that bind: Measuring the strength of consumers' emotional attachments to brands. *J. Consumer Psychol.* 15(1):77–91.

van Heerde HJ, Bijmolt THA (2005) Decomposing the promotional revenue bump for loyalty program revenue members versus non-members. *J. Marketing Res.* 42(4):443–457.