Can Non-tiered Customer Loyalty Programs Be Profitable?

Arun Gopalakrishnan

Zhenling Jiang

Yulia Nevskaya

Raphael Thomadsen[1]

[1] Corresponding Author. Olin Business School, Washington University in St. Louis, One Brookings Dr., Knight Hall 404, St. Louis, MO 63132. thomadsen@wustl.edu. 314-935-3573.

Can Non-tiered Customer Loyalty Programs Be Profitable?

Abstract

We study the impact of launching a non-tiered customer loyalty program on consumers' spending per visit, frequency of visits and attrition rates, as well as the overall customer value. We demonstrate these results both through descriptive difference-in-difference regressions as well as through a duration-dependent Hidden Markov Model we develop. We find 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 over 80% of the program's total lift, while 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.

1. Introduction

Customer loyalty programs are used by a wide range of businesses. In 2013, the average US 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 creates economic lock-in due to increasing benefits and consumer self-signaling (e.g., Drèze and Nunes, 2009; Orhun and Guo, 2019). By contrast, non-tiered programs, such as "buy 10 get 1 free," or "$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. Further, 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, 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%.[2]

[2] Taylor and Neslin (2005) also find that a similar program that rewarded 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.

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 (e.g., Bolton et al., 2000; Lal and Bell, 2003) have shown sizeable 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 dataset 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.[3]

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). While 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 utilize a number of novel aspects in our data and estimation approach. First, we collect data for a cohort of over 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 the 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[4] (which we refer to as "non-members").

[3] We use the term Customer Lifetime Value (CLV) to denote a five-year time horizon in our data setting, as explained Section 4.2.
[4] 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 prior to the program results in very similar pre-program behaviors.

On the methodology side, we bring together the evaluation of non-tiered loyalty programs and the literature on customer-lifetime-value models (e.g., 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) (e.g., Netzer et al. 2008) that integrates customer attrition and frequency of visits. This model allows for the estimation of the program's overall effectiveness and the 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 the program introduction for automatic members and non-members, and attribute the difference in the change between the automatic and non-members to the effect of the program (i.e., a difference-in-differences approach).

The rest of the paper is organized as follows. In section 2, we describe our data, 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. Dataset and Descriptive Analysis

In this section, we first describe our data in subsection 2.1. We then provide some descriptive analyses in subsection 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 dataset 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.[5] All

[5] These customers do not have a visit during the six-month window prior to their observed first visit in our dataset. Hence, we assume these customers are newly acquired during their first observed visit.

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.[6]

In our dataset, the price of a haircut increased by $1 a 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

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

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. Due to this self-selection, the behavior of a particular individual before and after they join the loyalty program may reflect higher usage 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 prior to 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 towards discounts, which could impact their response to the loyalty program. However, no information is available in our dataset 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. Note that only these redeeming members entail costs to the company from running the loyalty program.[7]

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

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

spending before accounting for applied discounts), days between visits, and the 6-month hiatus rate, which we define for the purposes of the descriptive analysis as an event where a customer does not visit for over 182 days after a purchase occasion. The threshold of 6 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 6 months between haircuts. Further, while we use this length of hiatus for our descriptive analysis, we **do not impose such a limit in our model-based analysis**. Compared to 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.

Table 1: Summary Statistics Customer Cohort Composition

Panel A: Customer Cohort Composition

	Number of Customers	Share of Customer Base	Number of Transactions mean (std. dev)	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

	Net Spending per Visit ($) mean (std. dev)	Gross Spending per Visit ($) mean (std. dev)	Days between Visits mean (std. dev)	6-Month Hiatus Rate mean (std. dev)
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%

2.2 Descriptive Analysis

In this subsection, we present some descriptive analysis about how the loyalty program affects different aspects of customer behavior: frequency of visits, spending per visit, and 6-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, as 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 the 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. Note that we do not include the 321 customers who made seven or more trips to the firm before the program introduction because these customers vary widely in the number of visits they made pre-program, and we have too few customers 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 the model-based analyses.

Table 2: Number of Automatic 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

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 1 if customer i is an automatic member and 0 if non-member, AP_{ik} is an indicator that equals 1 if customer i visit to k occurs after the program introduction, W_k is a vector of year-week dummies that represent a (non-parametric) 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 the program introduction, and ϵ_{ik} is an error term. The main parameter of interest, β_{ap}, is a difference-in-differences type measure that is identified by how much Y changes for automatic members after the program introduction compared to 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 defined as above. In this specification, the individual fixed effects capture the 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 to non-members.

The third strategy is to use propensity score matching. The goal is to find non-members that closely resemble the behavior of automatic members before the program introduction to serve as the control group. We use the 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.[8] 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 1 for automatic members and 0 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 to matched non-members after program introduction.

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 the introduction of the loyalty program, using the spending per visit of non-members over time as a non-parametric control for time trends. We use each of the three strategies described above to obtain a difference-in-differences measure of the program impact on spending for members.

Table 3. Spending per Visit

	Dependent variable:					
	Net Spending Per Visit			Gross Spending Per Visit		
	Segment FE (1)	Individual FE (2)	Propensity Score Matching (3)	Segment FE (4)	Individual FE (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 Visit Before Program Dummies	Yes	No	No	Yes	No	No
Number of Visit 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

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

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
Note:						$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

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 after accounting for any rewards coupon they 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 they 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.

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 that return within a year. Customers are most likely to revisit between 3 to 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). Note that 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) to (3), with Y_{ik} representing the number of days between the k^{th} 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.

Figure 1. Histogram of Number of Days between Visits

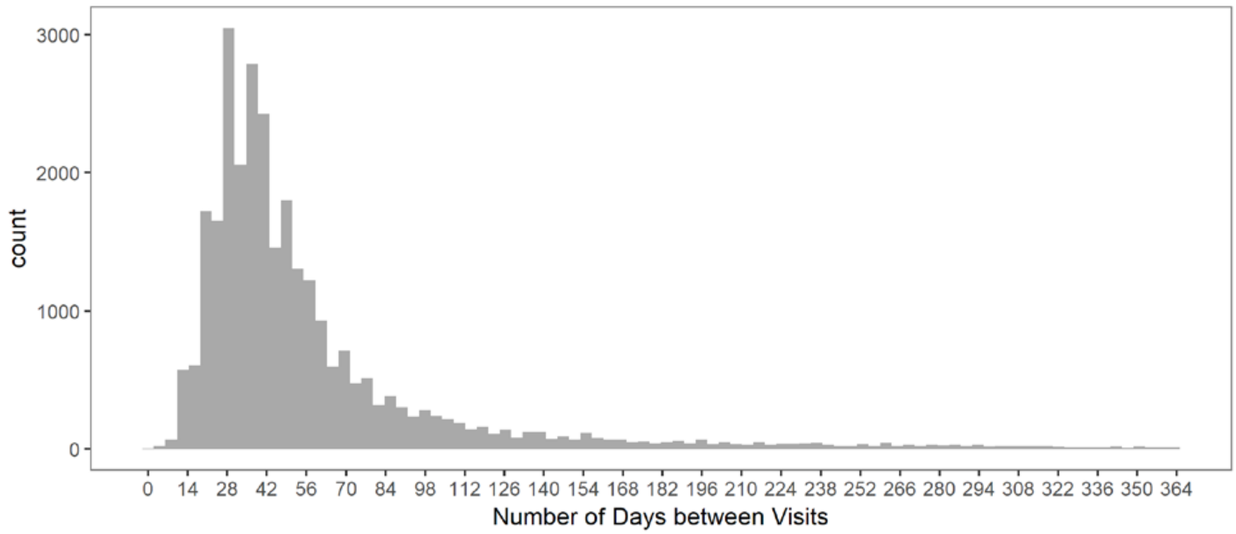


Table 4. Program Effect on Frequency (Days between Visits) and 6-month Hiatus

	Dependent variable:				
	Days between Visits (OLS)			6-month Hiatus (Logistic regression)	
	Segment FE (1)	Individual FE (2)	Propensity Score Matching (3)	Segment FE (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 Visit Before Program Dummies	Yes	No	No	Yes	No
Number of Visit 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
Note:					*p<0.1; **p<0.05; ***p<0.01

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 (6 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 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 6-month hiatus is not a repeated event for most customers. We therefore use the matched segments with the same number of visits as well as 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 1 if customer i does not revisit within 182 days after visit k, and 0 otherwise.

The results are shown in columns 4 and 5 of Table 4. After controlling for the time trend by the year-week fixed effects, we see that automatic members do indeed have a lower probability of hiatus after the 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 the 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 (e.g., 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 subsection 3.1. In subsection 3.2, we lay out the model framework. In subsection 3.3, we discuss identification. In subsection 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 program's impact on the frequency of visit 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 4 to 6 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 mis-estimation 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 inter-visit duration with the *implicit* assumption that there is no attrition. In other words, a standard model is likely to be mis-specified 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. A Hidden Markov Model or HMM (e.g., 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 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 subsection 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 1 if customer *i* visited the salon at time *t*, and 0 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: $p(Y_{it} = 1 | d_{it}, s_{it} = A) > 0$. On the other hand, the customer will not visit the salon in the dormant state, i.e., $p(Y_{it} = 1 | s_{it} = D) = 0$. 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:

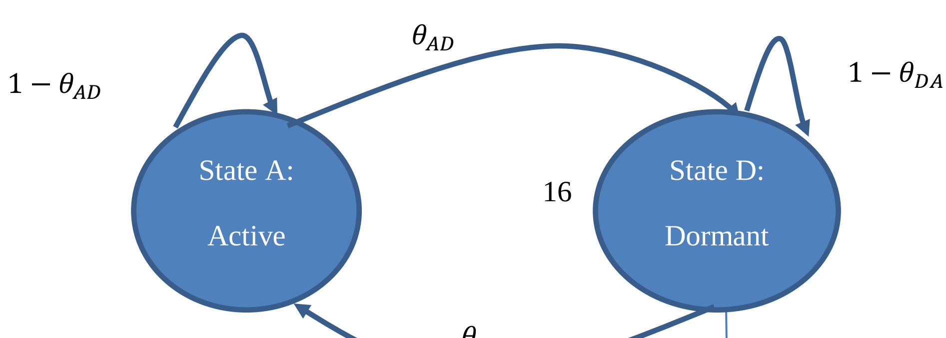
$$h_{it}(d_{it}) = p(Y_{it} = 1 | d_{it}, s_{it} = A),$$

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, i.e. $s_{it} = A$, as follows:

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

The customer also has some probability of transitioning between states A and D. The probability of transitioning from state A to 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 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 the dormant state to only occur immediately after a purchase for reasons that are explained in Section 3.3.

Figure 2: Graphical Representation of the Duration-Dependent HMM (DD-HMM)



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 1 if customer *i* chooses to visit at time *t*. As noted before, the initial state distribution is $p(s_{i0} = A) = 1$ because we assume the customer enters the panel data with their 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 $\boldsymbol{s_i^t} \equiv \{\mathrm{s}_{\mathrm{i}1}, \dots s_{it}\}$ 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

$$L\left(\theta_{AD}, \theta_{DA}, h_{it}(d_{it})\middle|\{Y_{it}\}, \boldsymbol{s_i^T}\right)$$

$$= \prod_{t=1}^{T} h_{it}\left(d_{it}(\boldsymbol{s_i^t})\right)^{I(Y_{it}=1, s_{it}=A)} \cdot \left[1 - h_{it}\left(d_{it}(\boldsymbol{s_i^t})\right)\right]^{I(Y_{it}=0, s_{it}=A)}$$

$$\cdot (\theta_{AD})^{I\left(s_{i,t-1}=A, s_{it}=D, Y_{i,t-1}=1\right)} \cdot (\theta_{DA})^{I\left(s_{i,t-1}=D, s_{it}=A\right)} \cdot (1-\theta_{DA})^{I\left(s_{i,t-1}=D, s_{it}=D\right)}$$

$$\cdot (1-\theta_{AD})^{I\left(s_{i,t-1}=A, s_{it}=A, Y_{i,t-1}=1\right)}. \tag{4}$$

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:

$$\ln(L(\theta_{AD}, \theta_{DA}, h_{it}(d_{it})|\{Y_{it}\})) = \sum_{q \in \boldsymbol{s_i^T}} \ln(L(\theta_{AD}, \theta_{DA}, h_{it}(d_{it})|\{Y_{it}\}, q)). \tag{5}$$

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

3.3 Model identification

In this subsection, we discuss why the general DD-HMM model defined in subsection 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 that 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 as well as 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 vs. dormant states could alternatively be explained through just a very flexible hazard function. Further, 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 non-parametric). 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 they last went grocery shopping, the more they are 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 non-parametric hazard function, however, can overfit the data even if it is restricted to be increasing. Note that a non-parametric hazard can have as many parameters as the maximum unique inter-visit duration in the data set. The number of data points for inter-visit 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 inter-visit durations to identify the parameters that best fit the data. Since 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, as shown below:

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

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 towards is p. With just three parameters (α, p, 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 $c > 1$) as this can be identified from the data (although we anticipate the increasing hazard given the logic above). We ultimately use this parametric hazard function because it avoids the overfitting concerns of the non-parametric 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 since a visit was observed. During weeks 1 through 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.

Table 5: Possible State Sequences for a Duration of Four Periods between Visits

State Sequence	Duration in active state	Intuition
AAAA	4 periods	Always active since last visit
DAAA	3 periods	Dormant for one period and returned to active
DDAA	2 periods	Dormant for two periods and returned to active
DDDA	1 period	Dormant for three periods and returned to active
ADAA	2 periods	Active for one period, then dormant for one period, and then returned to "restart" being active for two periods
AADA	1 period	Active for two periods, then dormant for one period, and then returned to "restart" being active for one period
ADDA	1 period	Active for one period, then dormant for two periods, and then returned to "restart" being active for one period
DADA	1 period	Toggling between dormant and active every time period, ending up in the active state

It is important to note that allowing for all of these possible state sequences (i.e., unrestricted state switching) leads to identification issues because the data does 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, they are 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, they 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, DDDA) would be permitted under this restriction whereas the last four would not (ADAA, AADA, ADDA, DADA). While we use the example of an inter-visit duration of 4 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, they stay there until the next visit, at which point they get 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 above state transition restriction is consistent with stylized customer behavior. We allow a customer to decide after a visit whether they want 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 inter-purchase time, as shown in Figure 1, has many shorter inter-purchase times, as we would expect given the periodicity of mens' haircuts. However, we also observe some very long inter-purchase times. These observed times would be highly improbable under a monotonic hazard function.[9] 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 inter-visit time being attributed to a defection to the dormant stage. Note that the threshold at which a customer is inferred as having attrited is probabilistic and different for each segment.

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

The 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 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 changes in θ_{AD}, the probability of attrition (transition from the active state to the dormant state), and in visit frequency (captured by hazard parameters, α, p, c, which we denote as a vector β for notational ease), after the loyalty program introduction,[10] we allow θ_{AD} and β to differ before and after the 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 $\theta_{AD}(LP_{it}) = \theta_{AD,before} \cdot I(LP_{it} = 0) + \theta_{AD,after} \cdot I(LP_{it} = 1)$, and $\beta(LP_{it}) = \beta_{before} \cdot I(LP_{it} = 0) + \beta_{after} \cdot I(LP_{it} = 1)$.

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.[11] 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).[12] We then estimate separate models for each segment of customers, as shown in Table 6. Note that we run

[10] We note the treatment is a regime change (launch of loyalty program) rather than a temporary benefit that expires.
[11] Examples can include opening of competing stores or changes in hair style preferences (e.g., keeping one's hair longer or shorter).
[12] The number of customers with seven or more visits before the program introduction was too low to reliably estimate treatment effects.

separate models for each segment,[13] as well as separate models based on whether the customers are automatic members or non-members. Thus, our difference-in-differences estimates are:

$$DID(\theta_{AD,s}) = (\theta_{AD,after,auto,s} - \theta_{AD,before,auto,s}) - (\theta_{AD,after,non,s} - \theta_{AD,before,non,s})$$

$$DID(h_{it,s}) = (h_{it,after,auto,s} - h_{it,before,auto,s}) - (h_{it,after,non,s} - h_{it,before,non,s}) \qquad (6)$$

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

Table 6: Estimated DD-HMM Models

Segment (Customer Type)	Automatic Members	Non-members
1 pre-program visit	DDHMM-1A	DDHMM-1N
2 pre-program visits	DDHMM-2A	DDHMM-2N
3 pre-program visits	DDHMM-3A	DDHMM-3N
4 pre-program visits	DDHMM-4A	DDHMM-4N
5 pre-program visits	DDHMM-5A	DDHMM-5N
6 pre-program visits	DDHMM-6A	DDHMM-6N

Each of the models in Table 6 has a total of 9 parameters (two sets of Discrete Weibull hazard parameters for before and after program introduction, and $\theta_{AD,before}$, $\theta_{AD,after}$, 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 dataset to capture the behaviors of the 6 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

[13] 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- and non- members.

after the 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 the after-program period. Further, 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 due to each of these conditions provides an indication of whether attrition or frequency effects contribute more towards changes in customer value.

4. Model Results

In this section, we present the results from the estimated model. In subsection 4.1, we show that the DD-HMM parameters for automatic members and non-members are generally similar before the program introduction. In subsection 4.2, we present the estimated effects of the program on attrition, frequency, and customer value. In subsection 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. While 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 the program introduction is reassuring. The full estimates of the DD-HMM model appears 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 6 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 while 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.

Table 7: Difference in DD-HMM Attrition and Hazard Parameters before Program Introduction

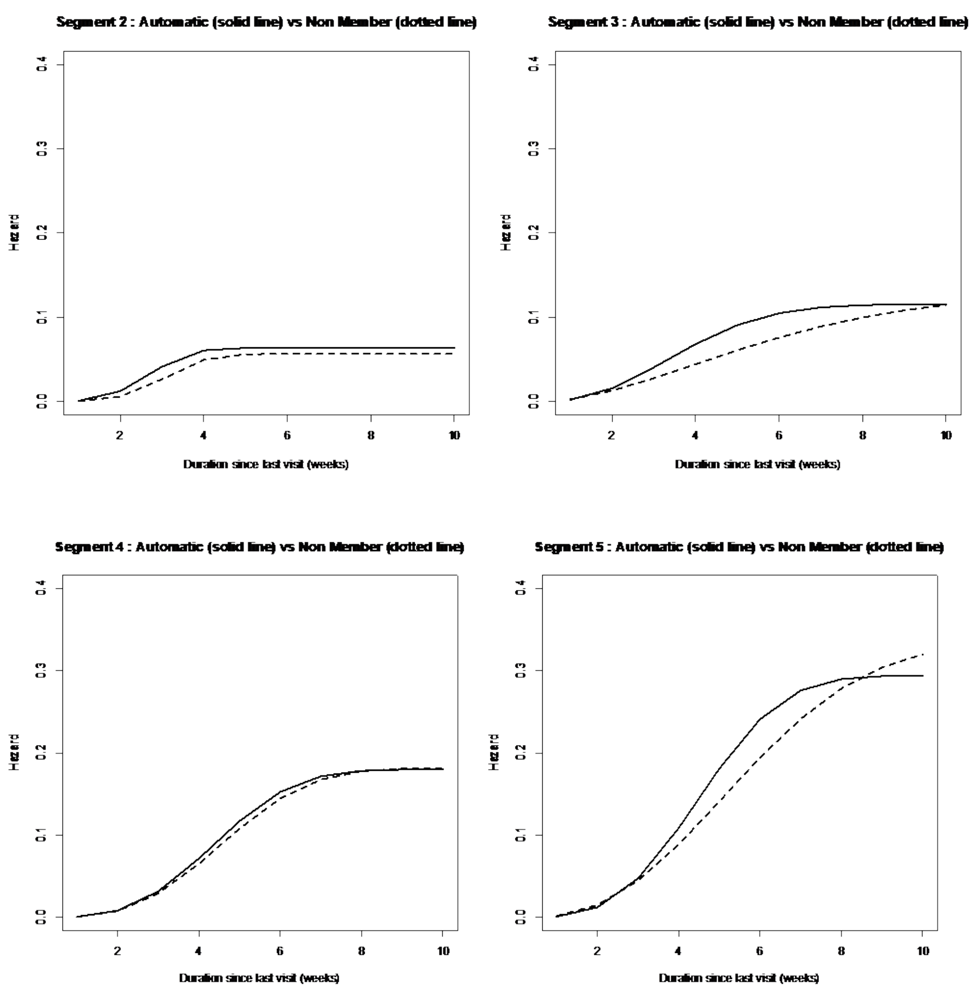
Variable	Segment (Pre-program visit frequency)					
	1	2	3	4	5	6
$\theta_{AD,before,auto} - \theta_{AD,before,non}$	-0.014	0.048	0.033	-0.000	-0.009	-0.003
$\alpha_{before,auto} - \alpha_{before,non}$	N/A	0.007	-0.002	-0.000	-0.002	**-0.002**
$p_{before,auto} - p_{before,non}$	N/A	0.007	-0.014	-0.002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

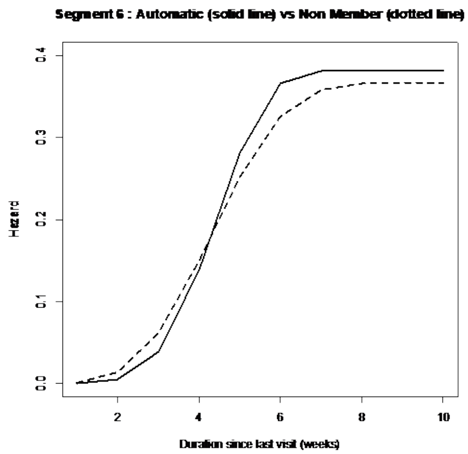
*Note: Differences that are statistically significant (p-value < 0.05) are **bolded**.*

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.

Figure 3: Pre-program hazard function for segments 2 through 6 comparing automatic- (solid line) and non- members (dotted lines)
Notes: Segment 1 has a zero hazard because the segment 1 customers do not have a repeat transaction.





We begin by discussing the direct difference-in-difference estimates of the program effect on the 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 analysis, 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 leads us to find that the loyalty program leads to a 23% (=5.4/23.3) relative reduction in the 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% to 19.5% in the attrition rate. The main reason for this difference lies in the more appropriate behavior-based definition of attrition used in our modeling approach as compared to the threshold of 182 days of inactivity used to define attrition in Section 2. Specifically, our model picks up attrition that occurs where 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 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 the program launch, we find that automatic members have an attrition rate that is 7.2 percentage points 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.

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
1 pre-program visit	-0.072	0.029	**0.013**
2 pre-program visits	-0.062	0.039	0.114
3 pre-program visits	-0.051	0.031	0.101
4 pre-program visits	0.001	0.023	0.717
5 pre-program visits	-0.030	0.024	0.219
6 pre-program visits	-0.052	0.020	**0.011**

Note: P-values less than 0.05 are in bold.

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 the 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%.[14] 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.

Table 9: Program Effect and Lift Estimates on Discounted Expected 5-Year Customer Value

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

[14] 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.

*Note: *** indicates p-value < 0.01,* ****** *indicates p-value < 0.05, * indicates p-value < 0.1*

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, due to the presence of the loyalty program. Finally, column 3 reports the percentage increase in the customer value as a result of the loyalty program. The program increases the customer value for all of the 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-value < 0.05) only for segments 1 and 6 in terms of revenue lift (both in 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, while providing a healthy return for highly frequent customers (i.e., segment 6) over a five-year horizon, are also effective at generating lift for the most infrequent customers.

The highly frequent segment 6 has the most economic incentive to stay with the firm as 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 6 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 these 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 Web 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 above, 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 the customer value can be directly attributed to reduced attrition. The results are shown in Table 10. They demonstrate a pattern similar to the one above, 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 Segment 1 and Segment 6. Note that 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.

Table 10: Attrition-Only Estimates on Discounted Expected 5-Year Customer Value

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

*Note: *** indicates p-value < 0.01,* ** *indicates p-value < 0.05, * indicates p-value < 0.1*

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 the average customer value changes when customers have the hazard rates of visits with the program compared to what their hazard rate would have been if the program had not been implemented. The results are in Table 11. We observe the aggregate lift due to frequency-only effects is 4.1% (compared to 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 to 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 inter-purchase period for many customers.

Table 11: Frequency-Only Estimates on Discounted Expected 5-Year Customer Value

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

*Note: *** indicates p-value < 0.01, ** indicates p-value < 0.05, * indicates p-value < 0.1*

In summary, we see the program significantly increases the 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 previous section. First, we present a DD-HMM model which 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 Web 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, each having its own transition probabilities and hazard functions. This extension could be useful if, within count segments, there are customers of different sub-types who may have differing behavioral responses to the loyalty program. We show in Web 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 experience about a 30% lift in value due to program introduction. This suggests that the count segments in Section 4.2 are relatively effective in capturing average program treatment effects.

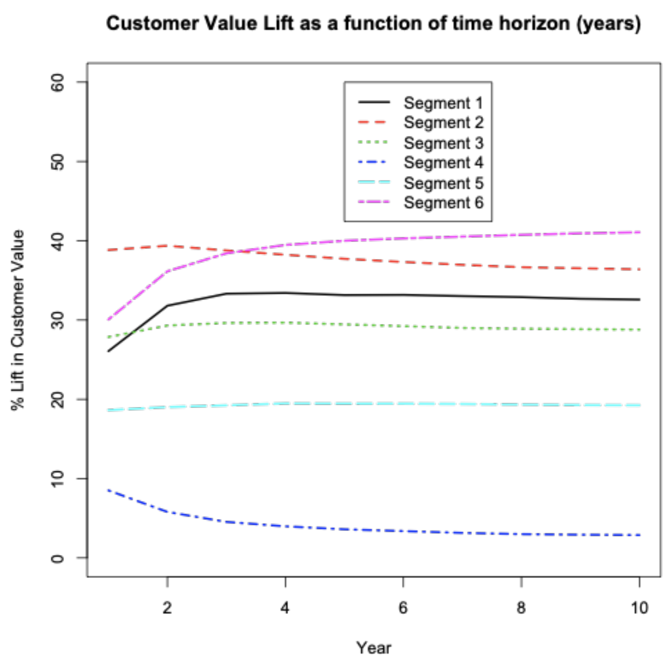
Our second robustness check concerns the sensitivity of our results to analyzing the CLV lift using a 5-year horizon on customer value. In Figure 4, we plot the percentage lift in customer value (a metric we report in Table 9 using a forward time horizon of five years) as a function of time horizon ranging from one to ten years. We note that while an infinite-horizon customer lifetime value 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.[15] 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.

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

Finally, Web 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 programs' value.

Figure 4: 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 the frequency and attrition, the finding that the magnitudes of the descriptive and modeled results are similar is reassuring because it suggests 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 the 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, while the customer that rarely comes in will have too low of 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. Note that this logic is driven by an up-sell and frequency mentality, and indeed we see that the moderate customers are the ones that 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, though consistent with previous literature, can be specific to the hair salon industry where the opportunities for upselling 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 Web Appendix B), could be consistent with economic incentives. On the other hand, only 2% of customers redeem a reward coupon. Further, 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 good will 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 (e.g., 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 a Hidden Markov Model, which 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 of 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 Hidden Markov Model can also be useful beyond the current application to measuring a loyalty program effectiveness as marketing managers can utilize the inferred knowledge about customer attrition for development of marketing tactics.

References

Ascarza, Eva, and Bruce GS Hardie (2013), "A joint model of usage and churn in contractual settings," *Marketing Science* 32(4), 570-590.

Berry, Jeff (2013). The 2013 COLLOQUY Loyalty Census. COLLOQUY Industry Report. http://www.colloquy.com/files/2013-COLLOQUY-Census-Talk-White-Paper.pdf

Bijmolt, Tammo H.A., Matilda Dorotic and Peter C. Verhoef (2011), "Loyalty Programs: Generalizations on Their Adoption, Effectiveness and Design," *Foundations and Trends in Marketing*, 5(4), 197-258.

Bolton, Ruth N., P. K. Kannan and Matthew D. Bramlett (2000), "Implications of Loyalty Program Membership and Service Experiences for Customer Retention and Value," *Journal of the Academy of Marketing Science*, 28(1), 95-108.

De Wulf, Kristof, Gaby Odekerken-Schröder, Dawn Iacobucci (2001), "Investments in Consumer Relationships: A Cross-Country and Cross-Industry Exploration," *Journal of Marketing*, 65(4), 33−50.

Drèze, Xavier and Joseph C Nunes (2009), "Feeling Superior: The Impact of Loyalty Program Structure on Consumers' Perceptions of Status," *Journal of Consumer Research*, 35(6), 890-905.

Fader, Peter S., Bruce GS Hardie, and Ka Lok Lee (2005), ""Counting your customers" the easy way: An alternative to the Pareto/NBD model," *Marketing Science* 24(2), 275-284.

Fader, Peter S. Bruce G. S. Hardie and Jen Shang (2010), "Customer-Base Analysis in a Discrete-Time Noncontractual Setting," *Marketing Science*, 29(6), 1086-1108.

Fader, Peter S., Bruce GS Hardie, Yuzhou Liu, Joseph Davin, and Thomas Steenburgh (2018), ""How to Project Customer Retention" Revisited: The Role of Duration Dependence." *Journal of Interactive Marketing*, 43, 1-16.

Fishbach, Ayelet, Rebecca K. Ratner, and Ying Zhang (2011), "Inherently loyal or easily bored?: Nonconscious activation of consistency versus variety-seeking behavior," *Journal of Consumer Psychology*, 21(1), 38-48.

Hartmann, Wesley R. and V. Brian Viard (2008), "Do frequency reward programs create switching costs? A dynamic structural analysis of demand in a reward program," *Quantitative Marketing and Economics*, 6(2), 109-137.

Henderson, Conor M., Joshua T. Beck and Robert W. Palmatier (2011), "Review of the theoretical underpinnings of loyalty programs," *Journal of Consumer Psychology*, 21(3), 256–276.

Kopalle, Praveen K., Yacheng Sun, Scott A. Neslin, Baohong Sun, Vanitha Swaminathan (2012), "The Joint Sales Impact of Frequency Reward and Customer Tier Components of Loyalty programs," *Marketing Science*, 31(2), 216-235.

Lacey, Russell, Jaebeom Suh, and Robert M. Morgan (2007), "Differential Effects of Preferential Treatment Levels on Relational Outcomes," *Journal of Service Research*, 9(3), 241–256.

Lal, Rajiv and David E. Bell (2003), "The Impact of Frequent Shopper Programs in Grocery Retailing," *Quantitative Marketing and Economics*, 1(2), 179-202.

Leenheer, Jorna, Harald J. van Heerde, Tammo H.A. Bijmoltc and Ale Smidtsd (2007), "Do loyalty programs really enhance behavioral loyalty? An empirical analysis accounting for self-selecting members," *International Journal of Research in Marketing*, 24(1), 31-47.

Lewis, Michael (2004), "The Influence of Loyalty Programs and Short-Term Promotions on Customer Retention," *Journal of Marketing Research*, 41(3), 281-292.

Liu, Yuping (2007), "The Long-Term Impact of Loyalty Programs on Consumer Purchase Behavior and Loyalty," *Journal of Marketing*, 71(4), 19-35.

Nakagawa, Toshio, and Shunji Osaki (1975), “The discrete Weibull distribution.” *IEEE transactions on reliability*, 24(5), 300-301.

Netzer, Oded, James M. Lattin, and V. Srinivasan (2008), "A hidden Markov model of customer relationship dynamics," *Marketing Science*, 27(2), 185-204.

Orhun, A. Yesim, and Tong Guo (2019), "Reaching for Gold: Frequent-Flyer Status Incentives and Moral Hazard," *working paper*.

Schmittlein, David C., Donald G. Morrison, and Richard Colombo (1987), "Counting your customers: Who-are they and what will they do next?" *Management Science* 33(1), 1-24.

Schwartz, Eric M., Eric T. Bradlow, and Peter S. Fader (2014), "Model selection using database characteristics: Developing a classification tree for longitudinal incidence data," *Marketing Science* 33(2), 188-205.

Schweidel, David A., Eric T. Bradlow, and Peter S. Fader (2011), "Portfolio dynamics for customers of a multiservice provider," *Management Science* 57(3), 471-486.

Sharp, Byron and Anne Sharp (1997), “Loyalty Programs and their impact on repeat-purchase loyalty patterns,” *International Journal of Research in Marketing*, 14(5), 473-486.

Stourm, Valeria, Eric T. Bradlow, and Peter S. Fader (2015), “Stockpiling Points in Linear Loyalty Programs,” *Journal of Marketing Research*, 52(2), 253-267.

Taylor, Gail Ayala and Scott A. Neslin (2005), “The Current and Future Sales Impact of a Retail Frequency Reward Program,” *Journal of Retailing*, 81(4), 293-305.

Thomson, Matthew, Deborah J. MacInnis, and C. Whan Park (2005), “The Ties That Bind: Measuring the Strength of Consumers’ Emotional Attachments to Brands,” *Journal of Consumer Psychology*, 15(1), 77-

91.

van Heerde, Harald J. and Tammo H.A. Bijmolt (2005), "Decomposing the Promotional Revenue Bump for Loyalty Program Revenue Members Versus Nonmembers," *Journal of Marketing Research*, 42(4), 443-457.

Appendix: Full Set of DD-HMM parameter estimates

Model Estimation Algorithm

We estimate the DD-HMM described in Section 3 of the paper 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 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 which 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 re-computed 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, while heterogeneity is flexibly captured by estimating separate models for each segment.

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.

$\theta_{AD}(LP_{it}) = \theta_{AD,before} \cdot I(LP_{it} = 0) + \theta_{AD,after} \cdot I(LP_{it} = 1)$, 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 as there is not enough variation in the data to estimate these before- and after- the program. Hence, a single θ_{DA} is estimated across the entire time horizon.

For the hazard parameters, we have 3 parameters under the discrete Weibull model for the before-program hazard and another 3 parameters for the hazard after the program is launched.

Estimated Parameters

We use Maximum Likelihood Estimation (MLE) to obtain parameter estimates for each DD-HMM. We present the full set of parameter estimates for the models for automatic (Table A1) and non- members (Table A2) below. Note that segment 1 does not have a pre-program hazard defined since it has no repeat visits after an initial transaction.

Table A1: DD-HMM Parameters for Automatic Members

Variable	Segment (Pre-program visit frequency)					
	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

Table A2: DD-HMM Parameters for Non-Members

Variable	Segment (Pre-program visit frequency)					
	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

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.

Web Companion Appendix to

Can Non-tiered Customer Loyalty Programs Be Profitable?

Web Appendix A: Robustness check using a heterogeneous DD-HMM for each segment

Our model results in the main text (Table 9) used a DD-HMM in which customers within the same segment (i.e., same count of pre-program visits) shared the same model parameters. While the count segmentation provides benefits in terms of being easy to interpret and also controls for opportunities a customer may have had to provide their email address prior to program launch, we now present a heterogeneous version of the DD-HMM that allows for unobserved heterogeneity via a "mixture of HMMs." That is, instead of a continuous mixing distribution, which both imposes a parametric assumption about the distribution of unobserved heterogeneity and substantially increases the computational burden (as MCMC would be required), we posit that a customer within a count segment comes from one of two DD-HMM sub-models with some probability of belonging to each mixture. In one sense, this is akin to a latent class analysis (Kamakura and Russell 1989) but now applied to an entire DD-HMM as a dynamic model within each class or mixture.

To estimate this model, the likelihood function shown in equation (5) is estimated for each latent class/mixture which has its own DD-HMM parameters, and then integrated over the uncertainty of which class a given customer belongs to (which is represented by a parameter). We then take the difference-in-differences approach outlined in Section 3.4 but accounting for this latent mixture approach. We allow for two latent mixtures in order to avoid overfitting issues. We reproduce the lift in customer value over a five year period using this heterogeneous DD-HMM approach in Table WA-1 to examine the impact of allowing for customers within a segment to have unobservable differences in their parameters.

Table WA-1: Program Effect and Lift Estimates on Discounted Expected 5-Year Customer Value using heterogeneous DD-HMM

Segment	Customer Value *Without* Program (Std. err.)	Change in Customer Value (Std. err.)	Lift in Customer Value (Std. err.)
1	\$62.22	\$30.50**	51.2%*
	(\$7.53)	(\$15.30)	(28.0%)
2	\$114.02	\$16.83	16.6%
	(\$15.59)	(\$19.39)	(19.6%)
3	\$134.54	\$39.70	31.6%
	(\$17.71)	(\$25.43)	(22.8%)
4	\$250.80	\$7.62	6.2%
	(\$34.95)	(\$47.27)	(29.4%)
5	\$345.48	\$5.85	3.0%
	(\$41.91)	(\$45.89)	(13.7%)
6	\$325.50	\$105.61**	33.7%*
	(\$39.17)	(\$52.66)	(18.4%)
Overall	\$138.46	\$28.64***	32.3%**
	(\$9.20)	(\$10.93)	(13.0%)

*Note: *** indicates p-value < 0.01,* ** *indicates p-value < 0.05, * indicates p-value < 0.1*

As can be seen by comparing Table WA-1 and Table 9 in the main text, the overall effect of the program remains around a 30% lift in customer value over 5 years, which is largely driven by improvements in segment 1 and 6 customers. This analysis demonstrates that Table 9 is reasonably robust to the addition of unobserved heterogeneity in the model.

Web Appendix B: Behavior Change around a Reward Visit

Prior literature suggests that consumers accelerate their purchases as they get closer to receiving a reward, a phenomenon known as "points pressure" (Kivetz et al., 2006; Hartmann and Viard, 2008; Kopalle et. al. 2012; Bijmolt et. al. 2011). We explore potential effects by examining *within individual* behavioral changes prior to receiving a reward and afterwards. Given that we cannot use individual fixed effects when studying attrition, we focus on how the spending and frequency of purchase changes in the reward cycle from 2 visits before receiving a reward to 2 visits after reward.[16] Consistent with prior literature, we find that consumers exhibit purchase acceleration – members accelerate the timing of the next visit as consumers move closer to getting their reward. The effect is the largest on the visit where they use a reward coupon.

	Dependent variable:	
	Net Spending Per Visit	Days between Visits
	(1)	(2)
2 Visits before Reward	-0.3894	-4.1128
	(0.6755)	(2.7555)
1 Visit before Reward	0.4563	-6.5210^{*}
	(0.7563)	(3.6162)
Reward Visits	-3.4345^{***}	-7.8857^{***}
	(0.9988)	(2.8305)
1 Visit after Reward	-0.2065	-1.7791
	(0.5333)	(3.4298)
2 Visits after Reward	-0.5502	-2.3093
	(0.6138)	(2.4629)
Program Start: Automatic Members	-0.1145	-2.1426
	(0.1735)	(1.4199)
Individual Dummies	Yes	Yes
Year+Week Dummies	Yes	Yes
Observations	21,995	15,578
Note:		$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

[16] Note that only members who have redeemed a reward during the data observation period will inform these parameter estimates.

Web Appendix C: Allowing for heterogeneous customer spend amounts in Customer Value analysis

In our main analysis, we assumed that each visit would incur a $21 monetary spend. In this appendix, we allow for a segment-specific distribution of monetary spends using the empirical mean and standard deviation in the data. In Table C-1, we show the monetary spend per visit broken down by segment. As can be seen, all segments have an average spend of around $21.

Table C-1: Monetary Spend by Segment

Pre-program Visit Frequency	Mean	Std Dev.
1	$20.61	$7.37
2	$20.60	$6.34
3	$20.91	$6.82
4	$20.79	$6.34
5	$21.17	$6.31
6	$21.20	$5.83

We now compute customer value lift (akin to Table 9) but allowing each visit to incur a spend amount from the segment-specific distribution above (we use a normal distribution with the corresponding mean and standard deviation) and show the results in Table C-2. The results are virtually unchanged as compared to Table 9, showing that the analysis is robust to allowing for variation in monetary amounts.

Table C-2: Program Effect and Lift Estimates on Discounted Expected 5-Year Customer Value using segment-specific monetary spend distributions

Segment	Customer Value *Without* Program (Std. err.)	Change in Customer Value (Std. err.)	Lift in Customer Value (Std. err.)
1	$71.22	$21.66**	31.6%**
	($7.62)	($10.90)	(16.0%)
2	$111.85	$40.15**	38.3%*
	($17.10)	($20.05)	(22.0%)
3	$144.76	$39.59	29.1%
	($18.32)	($25.43)	(20.1%)
4	$247.68	$6.13	4.6%
	($38.72)	($42.92)	(18.0%)
5	$303.75	$51.32	19.0%
	($44.50)	($48.97)	(17.8%)
6	$319.92	$118.71**	39.1%**
	($45.13)	($52.66)	(19.9%)
Overall	$139.52	$33.28***	28.9%***
	($7.85)	($9.03)	(8.6%)

*Note: *** indicates p-value < 0.01, ** indicates p-value < 0.05, * indicates p-value < 0.1*