Essays on Econometric Marketing: Three Empirical Studies on Customer Purchase Behavior

Kazuki Nishio

Graduate School of Economics, Keio University

Contents

1 Overview of the Dissertation

Customer purchase behavior, which has been quantitatively investigated in the literature on marketing science and econometric marketing, is difficult to capture and predict precisely. This is due to customers' incidental shopping behavior, heterogeneity among customers, and various marketing strategies implemented by stores and companies. To date, various econometric models have been developed and applied to customer purchase data to understand the dynamics of customer purchasing activity. However, there is still room to apply other econometric models to the customer purchase data for more accurate estimation and prediction.

Thus, this dissertation proposes econometric models that have not been focused on in the field of marketing science. The dissertation consists of three empirical studies and designs the model structure in each study while considering the characteristics of the targeted marketing phenomenon. The first study applies a marked point process model to the purchasing data of customers who participate in a customer tier program (a type of loyalty program). The second study employs the Weibull proportional hazards model, the generalized linear model, and the Pareto/NBD and gamma-gamma models to estimate the influence of promotional incentives in a loyalty program on the points pressure effect and customer lifetime value. The third study applies a marked Hawkes process to the customer purchase data in retail stores to accurately capture the joint dynamics of customer purchase frequency and amount. The following are the abstracts of each study in this dissertation:

Study 1: Customer Tier Programs and Marked Point Process Models (Chapter 2)

The first study in this dissertation investigates the dynamics of customer purchase behavior within a customer tier program, with a particular focus on the points pressure effect. This effect is derived from the goal gradient hypothesis, which suggests that organisms, whether humans or animals, accelerate their actions as they near a goal. In the context of customer tier programs, this manifests as increased purchase frequency and higher spending per store visit as customers approach a tier threshold that offers enhanced rewards or valuable services in the next period.

This study examines customer purchase data from a large Japanese supermarket chain that operates a customer tier program. The program assigns customers to different tiers based on their total spending in the last month, with each tier offering progressively better rewards. Customers who spend over 100 (200) dollars within a given month qualify for the "gold" ("platinum") tier in the following month. Each tier increases the rate at which customers earn points. Specifically, customers in the gold (platinum) tier can enjoy 1.5 (2) times higher point return rate than the ones in the "normal" tier.

To model the interaction between the timing of purchases and the amounts spent, this study employs a marked point process model that has often been used in the context of seismology, criminology, and finance. The marked point process model provides two key advantages. First, it allows for the modeling of purchase events as probabilistic occurrences, which better reflects real-world customer purchase behavior. Second, it captures the interaction between purchase intervals and amounts, offering a more holistic view of how customers respond to customer tier programs.

The results of the study show that the points pressure effect is particularly strong among customers who originally visited the store infrequently and had not attained the threshold established in the customer tier program many times in the past. In other words, originally loyal customers are less likely to be influenced by the points pressure. In addition, this study demonstrates that the points pressure acting on the way to the first threshold is stronger than the one to the second threshold, and the customer tier program has a stronger positive effect on purchase frequency than on purchase amount. Furthermore, it reveals that the three-tier customer tier program is superior to the two-tier one with respect to the operating income in supermarkets. Overall, this study contributes to the literature on loyalty programs by revealing a profitable source of the customer tier program.

Study 2: Effects of Promotional Incentives in a Loyalty Program (Chapter 3)

The second study in this dissertation examines the short- and long-term effects of promotional incentives within a loyalty program on customer purchase behavior. Specifically, the study investigates the influence of these incentives on the points pressure effect (the short-term behavioral phenomenon where customers accelerate their purchases as they approach a reward threshold) and customer lifetime value (the estimated present value of the net benefit a customer provides to the firm).

This study uses ID-POS data from a major Japanese retail company that operates a loyalty program incorporating birthday rewards as promotional incentives. Customers receive additional benefits during their birthday month, such as a double points return rate and a monetary reward (in the form of shopping points) after making a purchase. These incentives are designed to encourage customers to visit the store and make purchases during this specific period. However, there has been a notable absence of studies investigating the actual effectiveness of such incentives from both short- and long-term perspectives.

To quantify the short- and long-term effects of promotional incentives in a loyalty program on customer purchase behavior in terms of the points pressure effect and customer lifetime value, this study employs the Weibull proportional hazards model, the generalized linear model, and the Pareto/NBD and gamma-gamma models after implementing propensity score matching. The former two models allow separate estimation of the strength of points pressure on customer purchase intervals and amounts by the reparameterization technique. On the other hand, the latter two models allow estimation of the lifetime value of each customer as the product of contribution margin ratio, revenue per transaction, and discounted expected transactions.

The results of the study indicate that customers become temporarily less influenced by the points pressure with respect to their purchase frequencies after receiving promotional incentives in a loyalty program. In addition, this study reveals that such incentives have no positive effect on increasing the customer lifetime value, which has rich managerial implications. Furthermore, it demonstrates that customers who do not exhibit active purchasing behavior for a specific period after impulsive buying triggered by these incentives have particularly low lifetime values. Overall, this study contributes to the literature on loyalty programs by revealing the profitability of promotional incentives within a loyalty program.

Study 3: Customer Purchasing Behavior and Marked Hawkes Processes (Chapter 4)

The third study in this dissertation investigates the applicability of marked Hawkes processes to ID-linked customer purchase data in retail stores. Since customer purchase intervals and amounts are interrelated, the marked Hawkes process that can simultaneously model the event occurrence times and its attributes (marks) can be appropriate for accurately estimating and predicting the overall dynamics of customer purchasing behavior.

The marked Hawkes process has primarily been used in the context of natural sciences, such as seismology, epidemiology, and meteorology. In addition to the simultaneous modeling described above, it offers several advantages for modeling. First, this model can take into account all the past event occurrence history without reducing the volume of usable data. Second, it allows for flexible modeling including the incorporation of random effects parameters, event-related covariates, and non-linear functions into the conditional intensity function. Third, the model allows for the flexible structuring of each sub-event occurrence.

Through the implementation of the EM algorithm, this study models the joint mechanism of customer purchase frequencies and amounts. In other words, the marked Hawkes process proposed in this study effectively identifies the factors influencing customers' current willingness to purchase, either positively or negatively. In addition, the estimation of random effects parameters leads to rich managerial implications for targeted marketing strategies. Furthermore, by simulating using the thinning method and comparing the proposed model's performance with a traditional model (log-linear regression model with the random intercept), this study shows that the proposed model greatly outperforms the traditional one in long-term forecasting of customer purchasing behavior. Overall, this study contributes to the literature on consumer behavior and point processes by demonstrating the usefulness and high predictive performance of the marked Hawkes process when applied to the ID-POS data in retail stores.

2 Empirical Research on Customer Tier Program Using the Marked Point Process [1]

2.1 Introduction

In today's era of information technology, retail stores are motivated to design a loyalty program based on the accumulated transaction data of their customers. To cite a famous and successful example, Starbucks has implemented an app-based loyalty program that offers free drinks on its customers' birthdays and vouchers based on their number of points, resulting in a 7% sales increase in 2019. As another example, Sephora, a French multinational chain of personal care and beauty shops, has conducted a loyalty program that offers a free shipping option and special coupons according to its customers' tier, boasting 17 million members in North America alone and achieving a 22% increase in cross-sell revenue. In fact, it has been pointed out that an appropriate loyalty program that considers the behavioral characteristics of each customer increases their loyalty to the store (Meyer-Waarden, 2015). We may find it easy to understand the overall trend in consumer purchase action. However, drafting managerially beneficial and effectual loyalty program strategies based on big data is usually complex since it is difficult to comprehend the heterogeneity in customer purchase behavior and response to loyalty programs. If we intend to absorb these complicated dispositions of customers, it is crucial to select appropriate model. Concretely, incorporation of dynamic structure of their visits and expenditures into statistical models has more sophisticated and elaborate managerial implications.

Hence, we adopt a marked point process approach to measure and gain a helpful insight into the effect of loyalty program (especially customer tier component) on consumer behavior. The prime feature of this approach is that it simultaneously considers purchase interval and the purchase amount per shopping opportunity. While previous studies have analyzed them separately (Bijwaard et al., 2006; Yamada and Sato, 2021), it is natural to think of them as interrelated, which suggests the validity of marked point process model. In Section 2.4, we compare the simulation result of this model with that of two separate linear regression models and assert the superiority of our model. Additionally, in the marked point process approach, the purchase time and quantity are regarded

[1] This chapter is based on the paper "Joint modeling of effects of customer tier program on customer purchase duration and purchase amount" written by Nishio and Hoshino in 2022. It was published in *Journal of Retailing and Consumer Services*, a peer-reviewed journal in the marketing field.

as probabilistic occurrences, enabling us to consider the realistic phenomena such as prolongation of time between purchases after a massive purchase or increase in purchase frequency during a specific campaign period. We describe it in detail in Section 2.3.

Marketing researchers argue about the extent to which loyalty programs affect consumers' perception and sense of worth toward commodities and their prices (Bridson et al., 2008; Kopalle et al., 2012). Some researchers assert that customers who participate in a loyalty program become willing to pay for a certain retailer and are more active in their behavior because of their satisfaction, habituation, and expectation of future rewards (Kim et al., 2001; Gandomi and Zolfaghari, 2013; Chhabra, 2017). Others argue that the magnitude of change in customer purchase behavior affected by loyalty programs can be limited when considering the endogeneity of program membership or focusing on discounters (Sharp and Sharp, 1997; Leenheera et al., 2007; Bombaij and Dekimpe, 2020). These controversies never end; however, at least we can say that the design of a loyalty program defines its usefulness and effectiveness.

While we referred to all the corporate reward programs as "loyalty programs," there are broadly two types of program styles: frequency reward and customer tier. To put it simply, the former is "buy x times, then receive xx." and the latter is "once you qualify for xx tier, you will be guaranteed valuable services for a certain period." (Blattberg et al., 2008). For instance, a stamp card at a coffee chain is classified as a frequency reward program and a meal delivery service, which provides customers with free shipping for the next month if they spend a certain amount of money on this servicing, is categorized as a customer tier program. In this study, we addressed the theme of usefulness and profitability of a customer tier program; therefore, we plan to delve deeper into this kind of program.

Customer tier programs have utilized a predominant strategy in corporate profit making. For example, travelers who are predisposed to stay at a certain hotel are more likely to attain elite status (high tier), resulting in their cognizance of switching cost and spending more (Tanford, 2013). In online fashion markets, customers with a futuristic and brand-oriented sense tend to form a valuable segment. Companies can efficiently make a big profit by giving close attention to them in administering customer tier programs (Morisada et al., 2019).

When designing a customer tier program, we have to consider the fact that consumers who initially have a low or moderate patronage level can be powerfully subject to customer tier programs since they can further enhance loyalty (Liu, 2007; Reinartz and Kumar, 2000). These seemingly irrelevant customers are frequently ignored in the establishment of a firm's marketing strategies because such customers tend to be considered unprofitable to the company because, for example, their purchase frequency is low. However, it has been pointed out that there is a difference between behavioral loyalty (patronage), represented by purchase frequency, and attitudinal loyalty, represented by emotional commitment, and that latent profitable customers are low in the former and high in the latter (Tanford and Baloglu, 2013; Radder et al., 2015). Thus, framing a customer tier program considering such obscure and easily overlooked components based on transaction data can be profoundly important to most companies. In this study, we consider interaction terms between customer loyalty variables and variables representing the effect of customer tier program. Based on that, we evaluate our program such that it does not ignore consumers who do not originally contribute much to the store.

What makes the customer tier program effective in the first place? The origin goes back to an experiment in the 1930s. In this trial, the locomotion rates of rats confined in a maze in a sequence of partitions were measured (Hull, 1932). The result manifests that rats accelerate as they get closer to the goal; this phenomenon is called "the goal gradient effect." This behavioral characteristic is common in the marketing domain also. For example, people exert more commitment to savings as they approach the savings goal presented by a banker (Cheema and Bagchi, 2011). Players in the online game market deepen their relationship with the game as they perceive the proximity of goals set in it (Teng, 2017).

In fact, this attribute makes the customer tier program beneficial and effective. The customer tier program usually establishes a threshold that assures customers, who exceed it, the right to enjoy privileged services in the next period. Customers accelerate their purchase rates as the threshold nears (this phenomenon is referred to as "points pressure effect"), resembling the Hull's example (Kivetz et al., 2006; Sharma and Verma, 2014).

Several previous studies have conducted surveys on points pressure effect (Taylor and Neslin, 2005; Dorotic et al., 2014; Nastasoiu et al., 2021). They employ a variety of methods and contribute significantly to the company management. However, there is room to incorporate the structure that reflects the reality of customer purchase behavior into the model used. Since the arrival time and payment are interrelated and intrinsically random, employing a marked point process based approach is highly reasonable and allows management executives to comprehend heterogeneity among their customers from a unique point of view.

The point process approach is primarily applied in the fields of seismology (Ogata, 1988; Ogata, 1998; Ogata, 2017), criminology (Mohler, 2013; Brantingham et al., 2021), finance (Demoulin et al., 2005; Embrechts et al., 2011; Yang et al., 2017; Herrera and Clements, 2018), electricity price (Christensen et al., 2012; Clements et al., 2015), and so on. More recently, this type of model has been utilized in the context of COVID-19 (Chiang et al., 2021; Koyama et al., 2021). In spite of general versatility and ubiquity of the point process model, the number of prior researches that use it as a means for drawing up marketing strategies is scarce (Xu et al., 2014). This may be because in the context of marketing, the self-excitation mechanism, wherein the occurrence of an event increases the probability of the following event incidence, is apparently non-existent. In fact, this self-excitation mechanism has a high affinity to the point process model. However, the clock time and money amount associated with a bunch of purchase actions in customers are essentially influenced by their recent purchase behavior. More specifically, the willingness to purchase temporarily decreases after mass purchasing against the backdrop of home inventory, which can be regarded as the opposite effect but essentially the equivalent structure of self-excitation (Hendel and Nevo, 2006). Hence, assessing the change in customer buying behavior instigated by a customer tier program through the marked point process approach is appropriate and splendid as it can accurately evaluate the effect of customer tier program by considering the above structure.

In our target customer tier program, two thresholds define three distinctive classes in the following month. Our program stipulates that consumers who made a purchase of 100 (200) dollars [2] or more (tax-excluded) per month can enjoy 50 (100)% higher point redemption rate in the next month (the original point redemption rate is 1 point per 2 dollars and 10 cents). From now on, we will expedientially call these two privileged customer groups "gold" and "platinum." We can say that comparing these two types of point pressure effects with different strength in the same program that arise from two separate thresholds makes our research intriguing. The other details of program structure are provided in Section 2.3.

The remainder of this chapter consists of four sections. In Section 2.2, we explain our key concepts while reviewing the literature and formulating some hypotheses. Then, we go into detail about the data and model to use in Section 2.3, and present a series of results and their interpretations in Section 2.4. Finally, in Section 2.5, we conclude the chapter by clarifying managerial implications

[2] We used an exchange rate of 100 yen to the dollar since the original paper was written in 2021.

and limitations of our investigation.

2.2 Literature Review and Hypotheses Development

2.2.1 Customer tier program

In the early 2000s, Zeithaml et al. (2001) proposed the "customer pyramid" concept, a precursor to the customer tier program. They point out that since the utility function of each customer is different, their potential profitability and what they want from a company are also different, and that a company's profit can be maximized by allocating scarce resources and providing various services appropriately to each customer group that is stratified by the customer pyramid. In fact, First Union, the sixth largest bank in the U.S. at the time, color-coded its customers and offered extra customer services to those who were colored green, which can be seen as a pioneering example of a modern customer tier program.

Since then, based on the customer pyramid concept described above, the customer tier program has become increasingly implemented by firms in various fields vis-à-vis profit-seeking activities. Various studies on the appropriate design and profitability of this program have been conducted. Palmeira et al. (2016) show that low-frequency customers who cannot hope to achieve a high customer tier are likely to have a negative feeling about the tier structure itself, communicating with them while demonstrating concrete rewards is more effective than directly referring to tiers in increasing their motivation. In an attempt to simultaneously capture both the effects of frequency reward and customer tier, Kopalle et al. (2012) report that compared to the service-oriented customer group, the price-oriented customer group tends to place more emphasis on the frequency reward program, while the customer tier program is likely to be focused on regardless of customer orientation. Tanford et al. (2011) conducted a study on the relationship between the customer tier program and switching costs, that is, the financial, physical, and psychological costs incurred by customers when switching to another brand. Focusing on a hotel loyalty program, they suggested that customers who belonged to a higher tier had more emotional attachment to their preferred hotel and were less affected by price changes, resulting in higher switching costs and lower likelihood of outflows. Additionally, in subsequent research, it has been revealed that the customer tier program varies the attitudinal and behavioral loyalty levels of the program members, and segmenting them accordingly can facilitate better customer understanding and profit enhancement (Tanford, 2013; Tanford and Malek, 2015).

2.2.2 Points pressure effect

Hull (1932) laid the foundation for the goal gradient effect through his experiment with rats. At the time of its discovery, this effect of accelerating behavior as the goal nears was thought to be true only for primitive organisms. However, in recent years, it has become clear that this phenomenon can also be seen in the context of retail loyalty programs. As an example, Kivetz et al. (2006) found that in a real cafe's reward program, purchases accelerated as the goal of free coffee redemption approached. The goal gradient effect provoked by expiration date, threshold, privileged rank, and so on, set up in the loyalty program is usually referred to as the "points pressure effect."

Liu (2007) performed an analysis with a loyalty program in a convenience store chain. He measured a difference in the effectiveness of the loyalty program according to customer initial store usage level. He found that the loyalty program had a stronger positive impact on both purchase frequency and amount of light and moderate customers than of heavy customers against a backdrop of room to amplify the degree of patronage to the store. Additionally, Nunes and Drèze (2011) surveyed a frequent-flier program. Their research found that travelers who had originally used airlines frequently were less likely to be affected by the points pressure effect of this program. This is because they tend to spontaneously self-learn how much effort it takes to raise their tiers and systematically accumulate miles (points). Based on these preceding researches, we hypothesize the following.

H1: Customers with originally frequent purchasing behavior are less susceptible to the points pressure effect.

Taylor and Neslin (2005) investigated the effect of "turkey reward program" conducted by a supermarket chain on customer purchase behavior. They quantified the strength of points pressure effect while considering the heterogeneity of individuals. This heterogeneity was measured based on responses to a questionnaire. The results demonstrated that the points pressure effect was particularly strong among customers who did not originally focus on the loyalty program. On the other hand, the effect of customers' prior experience with a loyalty program on the points pressure effect has not been well investigated so far. In this study, we regard the number of customers' past loyal achievements as the degree of their attention to the loyalty program and then formulate the following hypothesis.

H2: The more times customers have reached the threshold in the past in the customer tier program, the less likely they are to be influenced by the point pressure.

Drèze and Nunes (2009) conducted a study on an optimal loyalty program design to maximize customer elitism. They note that three-tier programs are more likely to make elite members realize their higher status relative to two-tier programs, thereby suggesting the superiority of the former. However, there is no previous study that provides a direct solution regarding whether the first or second threshold has a stronger points pressure effect, and whether three-tier programs have a greater impact on purchase frequency or purchase amount. Mercedes and Óscar (2013) point out that customers who are loyal to a store tend to make planned purchases rather than impulsive ones. Using a game-theoretic model, Bazargana et al. (2017) show that in a competitive market, companies that establish reward expiry are more profitable than those that do not, as a backdrop of switching costs. Based on these two preceding studies, we can hypothesize that the first threshold has a stronger points pressure effect because customers with the ability to reach the second threshold are likely to be highly loyal and less likely to feel pressure, and the customer tier program has a greater impact on purchase frequency because the thresholds with a time limit set in the customer tier program are expected to discourage customers from impulsively making additional purchases at competing retail stores just before the time limit. We summarize the above-mentioned discussion as H3.

H3: The point pressure acting on the way to the first threshold (100 dollars in our customer tier program) is stronger than the one to the second threshold (200 dollars), and the customer tier program has a stronger positive effect on purchase frequency than on purchase amount.

2.3 Data and model description

2.3.1 Data

We use the point of sales data offered by TRIAL, a nationwide supermarket in Japan that has approximately 260 branches as of July 2021. Most of the stores are located in the suburbs and mainly deal with groceries and daily necessities. This company actively works on mapping out strategies based on the accumulated transaction data. The collected data files are separated on a monthly basis; they record customer card ID, per-second transaction time, transaction amount, amount of points awarded and used, accumulated point value, and kinds of transactions. Each file contains about 10 million transaction logs by nearly 2 million consumers. The purchase log of a specific customer can be traced using the card ID number. In this company, the customer tier program was executed from November 2013 to October 2017. Note here that in November 2017, this program transitioned to a frequency reward program where 1 point was awarded for every 2 dollars and 10 cents, including tax, and was available at 1 cent. This is beyond the scope of this study. Our customer tier program had a two-staged mechanism to accelerate customer purchase behavior, which makes our data unique and interesting. It is notable that the point usage system in which consumers can convert 1 point to 1 cent remained unchanged in this period. Additionally, since this company had adopted Everyday Low Price strategy and few special discounts were made during the period, it is possible to assume that the discount effect does not need to be considered or is at best absorbed in the day-of-week effect.

We perform two data processing operations prior to the analysis phase. First, we randomly choose 3000 customers who purchase at least once a month by using a simple random sampling method to resolve the censoring problem. The censoring can be caused by moving or switching to a different supermarket chain, and ignoring it diminishes the certainty of our estimation. Note here that the customers targeted in the analysis in this study are none other than shoppers at the supermarket, and they are randomly chosen from all the data across approximately 260 branches. In addition, the data used in this study capture approximately 95% of all purchases. More specifically, of the 3.3 billion dollars in total sales in 2015, approximately 3.1 billion dollars was associated with the use of point cards. Owing to computational problems, we use the data of only 3000 customers out of approximately 2 million, but because the percentage of point card users is very high, the data we use are sufficient to capture the overall trend in TRIAL. Second, we use the transaction data from January 2015 to December 2015 to calculate covariates of consumers, which enables our model to capture the effect provoked by the heterogeneity among customers. More specifically, the monthly average purchase amount, the average number of monthly store visits, and the total number of gold and platinum tiers achievement in this time span are computed, each representing the degree of customer loyalty. A note of caution is that the model parameters are estimated based on the one-year data from January 2016 to December 2016, not using the period for the covariate calculation.

In addition to the customer covariates, we consider more direct impacts on the customer purchase behavior, including variables that correspond to points pressure effect. Compared to the aforementioned covariates, these elements are time-dependent. More concretely, day of week of the transaction date, dummy variables which signify whether a customer belongs to the gold or platinum tier at a specific point of time, remaining amount up to the thresholds (100 and 200 dollars), elapsed time since the last purchase action, and previous purchase amount are considered in our model framework. Furthermore, to measure the strength of customer tier program impact on each customer group, interaction terms between customer covariates and variables indicating the effect of the program are also included in our model.

2.3.2 Model

The model in this study is based on the approaches adopted by Zhuang et al. (2002) and Mohler et al. (2011). Although the former is about earthquake occurrence and the latter is about incidence of crime, we dare to apply this methodology to the point of sales data on the premise of disputation presented in Section 2.1. The model based on point process is usually characterized by its intensity function. Here, the intensity specifies the instantaneous likelihood of an event happening. More specifically, if an event is likely to occur at a certain point of time, then the intensity at this point will take a large value. Now, we show the form of conditional intensity function used in this study.

$$\begin{aligned}\lambda_c\left(t, m \mid H_t\right) = \xi\nu\left(t\right)\mu\left(m\right) + \exp\{&\alpha_0 + \alpha_1\left(t - t_{c-}\right)\\ &+ \alpha_2\left(t - t_{c-}\right)^2 + \alpha_3 m_{c-} + \boldsymbol{x}_c^{\mathsf{T}}\boldsymbol{\beta} + \boldsymbol{x}_{t,c}^{\mathsf{T}}\boldsymbol{\gamma}\}\end{aligned}$$

There are five caveats in interpreting the meaning of the above formula. First, this expression is a "conditional" intensity function, indicating the momentary strength of a buying motive in consideration of purchase history. For example, customers who have not yet shown their purchase behavior near the end of the year tend to have a high conditional intensity against a backdrop of stockpiling. In the next place, the constant part $\xi\nu\left(t\right)\mu\left(m\right)$ depends on clock time and amount of payment, which enhances our model's explanatory power. If this term is constant in the literal sense, we cannot capture the heterogeneity associated with time of day and purchase amount. This term corresponds to the background intensity in the context of point process, which defines the intrinsic event occurrence rate such as strenuous shopper movement at the end of the year. Third, the exponential part contains elapsed time from the last purchase, previous purchase amount, time-independent loyalty variables, and time-dependent covariate vectors, which play an important role in compassing the heterogenous consumer purchase actions. Fourth, the above conditional intensity function is formularized for each customer as plural sequences of buying history in the customer group are assumed to be independent of each other. In other words, we assume that a series of purchase actions performed by a certain customer is not impacted by the existence of any other customer. Finally, the parameters are common within a customer group and do not depend on the customer index c.

Now, we summarize the definition of variables in Table 1. The variable names do not appear explicitly on the preceding formula, and the corresponding parameters (shown in the second column of Table 1) can be useful in grasping the gist.

The Maximum Likelihood Estimation is useful for estimating parameters, so we now project the form of log likelihood function.

$$l\left(\Theta\right) = \sum_{c=1}^{n_c} \sum_{i=l_c}^{l'_c} \log \lambda_c \left(t_i, m_i \mid H_{t_i}\right) - \sum_{c=1}^{n_c} \int_0^T \sum_{m \in \mathcal{M}} \lambda_c \left(s, m \mid H_s\right) ds$$

This functional form is commonly found in the context of point process with a few minor modifications such as the decomposition of joint probability density function into the product of marginal probability density functions, the details of which are mentioned in Daley and Vere-Jones (2003), for example. The notation is as follows: Θ is a parameter vector that consists of ξ, α_0, α_1, α_2, α_3, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$; n_c is the number of customers; l_c (l'_c) is the first (last) index in the purchase data of customer c; and $\mathcal{M}$ is the whole set of possible purchase amounts. Moreover, $t_{i,c-}$ ($m_{i,c-}$) is the most recent purchase time (amount) of customer c at the clock time t_i, though it does not emerge directly on the foregoing equations.

However, we cannot estimate parameters by simply trying to maximize the previous log likelihood function due to the existence of $\nu\left(t\right)$ and $\mu\left(m\right)$, which is a big problem. Therefore, to put it precisely, we must execute an iterative algorithm with variable bandwidth kernel density estimation. The details of this algorithm are provided in Subsection 2.6.1.

Table 1: Definition of variables

Variable name	Params	Definition
diffTime	α_1, α_2	Elapsed time since the last purchase of customer c.
previousAmount	α_3	The last purchase amount of customer c.
pastAmount	β_1	Monthly average purchase amount of customer c from Jan 2015 to Dec 2015.
pastRealVisit	β_2	Average number of monthly store visits by customer c from Jan 2015 to Dec 2015.
pastGoldCount	β_3	Total number of gold tier achievement by customer c from Jan 2015 to Dec 2015.
pastPlatinumCount	β_4	Total number of platinum tier achievement by customer c from Jan 2015 to Dec 2015.
MonFlag - SunFlag	$\gamma_1 - \gamma_6$	Day of the week flag of the transaction date (except for Thursday).
thisGoldFlag	γ_7	Dummy variable about whether customer c belongs to the gold tier in this month.
thisPlatinumFlag	γ_8	Dummy variable about whether customer c belongs to the platinum tier in this month.
before100	γ_9	Remaining amount up to 100 dollars.
before200	γ_{10}	Remaining amount up to 200 dollars.
interection1 - 16	$\gamma_{11} - \gamma_{26}$	Interaction terms between $\beta_1 - \beta_4$ and $\gamma_7 - \gamma_{10}$ (Here, $\gamma_{11} = \beta_1 \times \gamma_7$, $\gamma_{12} = \beta_1 \times \gamma_8$, $\ldots$, $\gamma_{26} = \beta_4 \times \gamma_{10}$).

2.4 Results

2.4.1 Summary statistics

Table 2 illustrates the summary statistics of variables. The number of customer visits in this study is 263,163. From the mean value and third quartile of pastGoldCount and pastPlatinumCount, it is considered that consumers tend to achieve the platinum tier beyond the gold tier in pursuit of higher redemption rate.

Table 2: Summary statistics of variables

Variable	Mean	Std	p25	p75
diffTime	4.058	4.682	1.001	5.819
previousAmount	25.056	23.165	9.690	32.940
pastAmount	227.167	170.762	113.007	295.098
pastRealVisit	10.362	8.195	5.333	12.833
pastGoldCount	3.775	3.187	1.0	6.0
pastPlatinumCount	4.903	4.687	0.0	10.0
before100	27.008	32.216	0.0	55.330
before200	15.191	29.159	0.0	9.230

Figure 1 shows the intensity distribution when the purchase actually happens. The mean value is 0.024 and the standard deviation is 0.019. Unlike the probability, the values of intensity themselves have no significant meaning; however, their ratio and magnitude relationship can be linked to the probability of event occurrence.

2.4.2 Main results

We now present our results and interpret them. Note that we standardized all the variables except for dummy ones beforehand to simplify the interpretation of coefficients. To begin with, we discuss the results for $\nu(t)$ and $\mu(m)$. These terms represent heterogeneity with respect to purchase time and amount, and play an important role in enhancing the model's explanatory power. However, as they are not directly related to the effects of the customer tier program that we focus on in this study, we will briefly discuss the results.

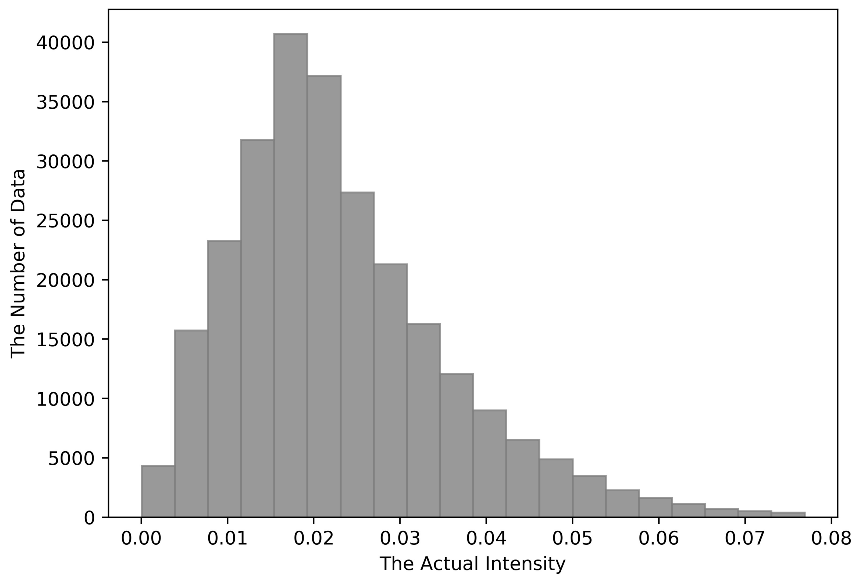

Figure 1: The intensity distribution

Figure 2 shows $\nu(t)$, with the horizontal axis being the follow-up period (from January to December 2016) and the vertical axis being the background intensity. To put it more clearly, at the time when $\nu(t)$ takes a large value, purchasing is more likely to occur for all customers in common, regardless of their attributes. From the figuration of this graph, precipitous increase in purchase demand is confirmed at the end of the year. This suggests that many customers tend to hoard ingredients for dishes served on New Year's Day. On the other hand, since few people bother to go to the supermarket during the New Year holidays, the value of $\nu(t)$ during this period is small. While the background intensity is roughly 0.008 at the beginning of the year, its value at the end of the year is approximately 0.015, which implies that purchase behavior is about twice as active during the year-end period compared to during New Year holidays. Additionally, a sharp drop in buying motivation can be observed in mid-August. This is because there is vacation of approximately 7 to 10 days in Japan around this period, and customers tend to stagnate their usual purchase actions against the backdrop of traveling, homecoming, and so on.

Figure 3 shows $\mu(m)$, with the horizontal axis being the purchase amount (up to 70 dollars) and vertical axis being the background intensity. More clearly, purchasing is more likely to occur

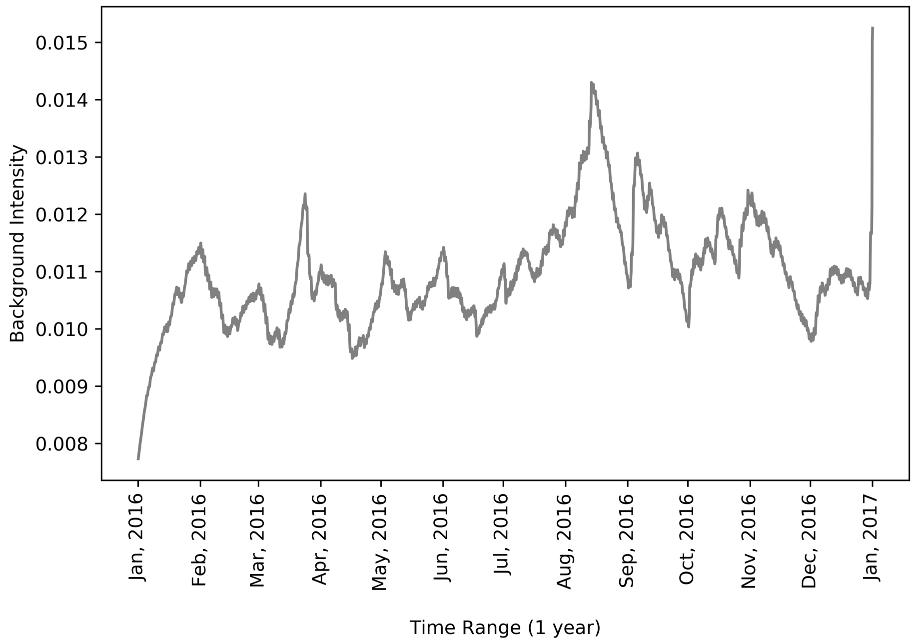

Figure 2: The background intensity with respect to time

in price ranges with high values of $\mu(m)$. The figure signalizes the effect of odd pricing policy that stimulates customers' willingness to purchase by setting item prices right before a nice round number such as 1 dollar 99 cents (Fraccaro et al., 2021). In other words, it suggests that many consumers are inclined to buy several products with odd pricing in a single purchase, resulting in the total purchase amount being just before a nice round number. Additionally, the background intensity is extremely low after 40 dollars. This is consistent with the third quartile of the previous amount in Table 2 and the fact that there are few high-value purchases at one time in supermarkets.

We will now discuss the results and interpretation of the parameter estimates. Table 3 shows estimated value (mean value), standard error, and confidence interval of each parameter. Here, we calculate these values from the results of applying our model to 1000 bootstrap samples. In the following, we intend to interpret parameters based on their mean values.

To begin with, the value of α_3 indicates that an increase of one standard deviation in the previous purchase amount decreases the intensity (the customer's willingness to purchase) by 11.54% [3]. Since only a limited amount of food can be stored at a time, if the last purchase amount is large,

[3] $\lambda_1 = \widehat{\xi}\overline{\nu}(t)\,\overline{\mu}(m) + \exp\{\widehat{\alpha}_0 + \widehat{\alpha}_1\left(\overline{t - t_{c-}}\right) + \widehat{\alpha}_2\left(\overline{t - t_{c-}}\right)^2 + \widehat{\alpha}_3\overline{m_{c-}} + \overline{\boldsymbol{x}_c}^{\top}\widehat{\boldsymbol{\beta}} + \overline{\boldsymbol{x}_{t,c}}^{\top}\widehat{\boldsymbol{\gamma}}\}$

Table 3: Estimation results

Params	Estimate	s.e.	C.I.
ξ	23.559	0.912	(21.740, 25.311)
α_0	-6.121	0.028	(-6.176, -6.068)
α_1	17.163	0.377	(16.443, 17.935)
α_2	-21.687	0.823	(-23.104, -19.757)
α_3	-0.263	0.011	(-0.285, -0.236)
β_1	-0.218	0.010	(-0.239, -0.194)
β_2	0.927	0.043	(0.873, 1.021)
β_3	-0.020	0.0006	(-0.021, -0.019)
β_4	-0.005	0.0001	(-0.005, -0.005)
γ_1	0.000	0.0001	(-0.0003, 0.0004)
γ_2	0.000	0.0001	(-0.0003, 0.0004)
γ_3	0.000	0.0001	(-0.0003, 0.0004)
γ_4	0.451	0.013	(0.426, 0.486)
γ_5	0.287	0.009	(0.266, 0.307)
γ_6	0.278	0.010	(0.254, 0.299)
γ_7	0.277	0.010	(0.257, 0.298)
γ_8	0.589	0.011	(0.567, 0.617)
γ_9	-1.071	0.016	(-1.103, -1.041)
γ_{10}	-0.178	0.004	(-0.187, -0.168)
γ_{11}	0.000	0.0001	(-0.0003, 0.0004)
γ_{12}	0.286	0.012	(0.260, 0.311)
γ_{13}	-0.060	0.002	(-0.065, -0.055)
γ_{14}	-0.010	0.0003	(-0.010, -0.009)
γ_{15}	-0.198	0.008	(-0.216, -0.180)
γ_{16}	-0.535	0.037	(-0.576, -0.443)
γ_{17}	0.132	0.005	(0.125, 0.147)
γ_{18}	0.050	0.002	(0.046, 0.053)
γ_{19}	0.030	0.001	(0.029, 0.033)
γ_{20}	0.030	0.001	(0.028, 0.032)
γ_{21}	0.143	0.005	(0.134, 0.155)
γ_{22}	0.000	0.0001	(-0.0003, 0.0004)
γ_{23}	0.062	0.002	(0.058, 0.067)
γ_{24}	0.210	0.009	(0.189, 0.229)
γ_{25}	0.329	0.010	(0.310, 0.355)
γ_{26}	-0.010	0.0003	(-0.011, -0.010)

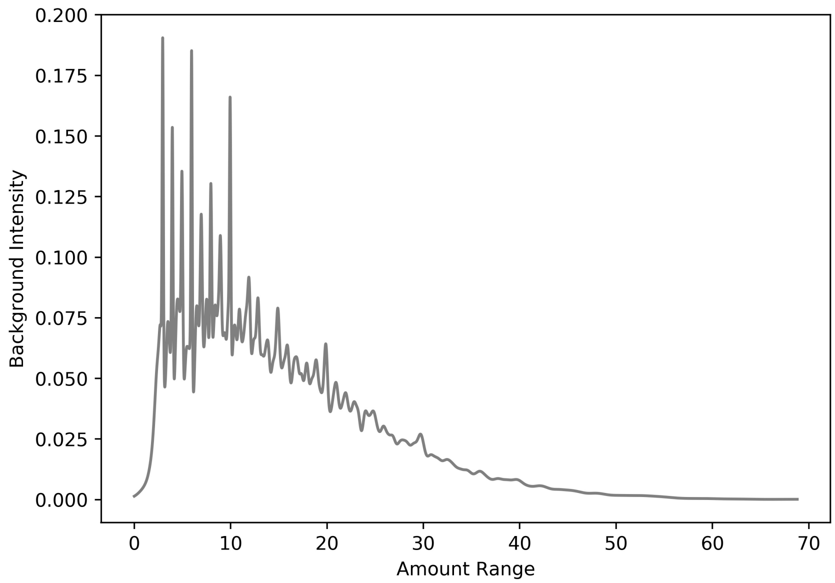

Figure 3: The background intensity with respect to money

the incentive for the next purchasing will be reduced.

Next, from the values of β_1, β_2, β_3, and β_4, we can say that to measure customer loyalty, the number of monthly store visits in the past is a more essential indicator than the monthly purchase amount and the number of annual gold/platinum tier achievement in the past. In fact, results indicate that an increment in the number of past monthly store visits by one standard deviation increases the intensity by 76.43%, while an increment in past monthly purchase amount by one standard deviation decreases the intensity by 9.77%. Additionally, the number of past annual gold/platinum tier achievement has almost no effect on the current intensity. The negativity of β_1 can be explained as follows. Of course, the past monthly purchase amount on its own indicates customer loyalty and positively influences the intensity, but since the other variable (corresponding to β_2) with which it is positively correlated (the correlation coefficient is 0.59) is rooted in the ease of store use, the effect on the intensity is dominated by this variable.

Moreover, the magnitude relationship between γ_1 through γ_6 demonstrates that the effect

$\lambda_2 = \widehat{\xi}\overline{\nu}\,(t)\,\overline{\mu}\,(m) + \exp\{\widehat{\alpha}_0 + \widehat{\alpha}_1\left(\overline{t - t_{c-}}\right) + \widehat{\alpha}_2\left(\overline{t - t_{c-}}\right)^2 + \widehat{\alpha}_3\left(\overline{m_{c-}} + (m_{c-})_{\mathrm{std}}\right) + \overline{\boldsymbol{x}_c}^{\mathsf{T}}\widehat{\boldsymbol{\beta}} + \overline{\boldsymbol{x}_{t,c}}^{\mathsf{T}}\widehat{\boldsymbol{\gamma}}\}$

$\rightarrow (\lambda_2 - \lambda_1)/\lambda_1 \approx -0.1154 \quad (\overline{x}: \text{mean value of } x)$

of the day of the week on the intensity is positively strong for Friday (most powerful), Saturday, and Sunday. For instance, the intensity goes up by 26.11% on Friday relative to on Thursday. In addition, the purchase motivation on Monday, Tuesday, and Wednesday is almost the same as that on Thursday. Here, it is signified that customers tend to purchase groceries for the following weekend on Friday and for the following weekday on Saturday and Sunday.

The most important areas to consider are those related to the values of γ_7 through γ_{26} because these values correspond to the effect of customer tier program and the points pressure effect. The values of γ_7 and γ_8 show that the changes in point redemption rate through the gold tier and the platinum tier increase the intensity by 15.26% and 34.63%, respectively. On another front, the values of γ_9 and γ_{10} clarify that as remaining amount up to the first (second) threshold decreases by one standard deviation, the intensity goes up by 32.62 (8.16)%. For example, the average time to a customer's next purchase with a monthly cumulative purchase of 50 dollars is approximately 30% longer than that of a customer with a monthly cumulative purchase of 80 dollars. Here, we must note that the pure points pressure effect that occurs when moving toward the first threshold is far stronger than the one when moving toward the second threshold, which supports the first half of H3. The above discussions do not consider the heterogeneity among customers, so we now focus on γ_{11} through γ_{26} to verify the effectiveness of our customer tier program in more detail. Among the estimated values of interaction term coefficients related to thisGoldFlag and thisPlatinumFlag, the values of γ_{15} and γ_{16} imply an interesting tendency. They show that an increase of one standard deviation in the number of past monthly store visits decreases the effect of gold (platinum) tier on the intensity by 17.95 ($=(1-\exp(\widehat{\gamma}_{15}))\times 100$) (41.43) %. This can be explained by the insight that customers who frequently visit the store tend to make a little amount of purchasing per visit, making it difficult for them to realize the change in point redemption rate. On the other hand, the values of γ_{17} and γ_{18} indicate that an increment in the number of past monthly store visits by one standard deviation decreases the points pressure effect acting on the way toward the first (second) threshold by 14.15 (5.10)%, which supports H1. The points pressure effect triggers additional store visits and purchases when cumulative purchase amount approaches the threshold, so this pressure is less likely to be applied to customers with a high level of usual store visits. Additionally, the values of γ_{21} and γ_{25} exhibit that an increment in the number of past annual gold (platinum) tier achievement by one standard deviation decreases the points pressure effect acting on the way toward the first threshold by 15.34 (38.95)%, which supports

H2.

2.4.3 Simulation and model comparison

In the previous subsection, we considered the estimates of background intensities and parameters. These considerations have significant implications; however, it is also important to provide a quantitative answer to how the existence of customer tier program affects customer purchase frequency and amount. Indeed, the previous results indicate that the effect of being platinum tier on the intensity is stronger than that of being gold tier against higher redemption rate ($\gamma_8 > \gamma_7$), while the points pressure with trying to achieve the platinum tier is weaker than the one with trying to achieve the gold tier since the accomplishment of platinum tier is often seen in customers who make purchases on a regular basis and they are less likely to be pressured ($-\gamma_{10} < -\gamma_9$). Here, the simulation becomes indispensable to answer the questions such as which threshold eventually has a stronger positive impact on purchase frequency and amount or whether there is a need for establishing three tiers in the first place.

Before we dive into the variable intervention through simulation, we demonstrate the superiority of our model to a traditional model that estimates purchase frequency and amount separately. Now, for the 3000 customers used in the previous estimation phase, we compare the true values of customer average of annual total visits (we call it AveVis from here), average purchase amount per visit (AveAmt), and annual total purchase amount (TotAmt) with the simulated values obtained from conventional linear regression model and our model.

Table 4: Model comparison

	AveVis	AveAmt	TotAmt
Actual Value	87.721	25.126	2204.078
Linear Regression Model	85.115	26.427	2249.334
Marked Point Process Model	87.550	24.828	2173.691

Table 4 shows that on AveVis and AveAmt, the linear regression model deviates from the true values by 2.97% and 5.18%, while our model deviates from them by only 0.19% and 1.19%. The difference in model performance derives from whether purchase frequency and amount are modeled simultaneously.

Now, we intervene interesting variables through the thinning method, the details of which are provided in Subsection 2.6.2. Specifically, the variables corresponding to γ_7 through γ_{10} are intervened to measure the effect of gold and platinum tiers on supermarket sales.

Table 5 shows the result of our simulation. Rows 2 to 7 in Table 5 respectively indicate the simulated result with a removal of the gold tier itself, effect of being the gold tier, points pressure effect acting on the way toward the first threshold, platinum tier itself, effect of being the platinum tier, and points pressure effect acting on the way toward the second threshold.

Table 5: Simulation result

	AveVis	AveAmt	TotAmt
Full Model	87.550	24.828	2173.691
Rm Gold	64.294	21.327	1371.198
Rm thisGold	85.035	24.508	2084.038
Rm PPtoGold	65.603	21.627	1418.796
Rm Platinum	71.124	22.667	1612.168
Rm thisPlatinum	76.064	23.364	1777.159
Rm PPtoPlatinum	80.928	23.998	1942.110

From the values of line 1, 2, and 5, we can say that a removal of the gold (platinum) tier itself leads to a decrease in AveVis, AveAmt, and TotAmt by 26.56 (18.76)%, 14.10 (8.70)%, and 36.92 (25.83)%, which indicates that our customer tier program affects purchase frequency about twice as much as purchase amount (supporting the latter half of H3). In addition, from the other lines, the gold tier influences supermarket sales mainly through the points pressure effect, while the platinum tier influences it mainly through a change in point redemption rate.

Based on the above simulation outcomes, we can judge whether the gold and platinum tiers are really needed. Our simulation is based on 3000 randomly chosen customers, but there are about 2 million customers in TRIAL, so we will consider the profit and cost for the entire company by simple multiplication in the following. Now, we assume that the profit margin of TRIAL is 1%. Thus, the increase in the company-wide annual profit due to the establishment of the gold tier becomes 16.05 million dollars. However, if we assume that the change in the points redemption rate through the gold and platinum tiers occurs according to the ratio of median values of pastGoldCount and pastPlatinumCount (both of which are 3) to 12, the company-wide annual increase in point values awarded due to the establishment of gold tier becomes 12.14 million dollars. Similarly, the company-wide annual increases in profit and cost through the installation of platinum tier are computed as 11.23 and 9.27 million dollars, respectively. Hence, we can conclude that the three-tier program is better than the two-tier program, although the increase in operating income is greater when the gold tier is established than when the platinum tier is established.

2.5 Concluding remarks

Our study investigated the influence of customer tier program with three tiers on customer purchase behavior. We found that there is a stronger pressure when proceeding to the first threshold than when proceeding to the second threshold, which can weaken according to the original customer loyalty level. Additionally, our research clarified the background intensity of purchase time and amount, customer loyalty variables that have a strongly positive impact on the current intensity, and quantitative impact of the gold and platinum tiers on supermarket sales and operating income. The innovative and remarkable aspect of our study was that the marked point process model enabled us to consider customer purchase time and amount simultaneously, and conduct flexible simulation with a high accuracy.

2.5.1 Research implications

This study has several research implications. First, the result of parameter estimates shows that supermarket managers should emphasize on the number of monthly store visits when looking for a customer group with potentially positive impact on current sales. A customer is likely to make a large purchase at a time in a specific month against a hosting of home party, extra bonus payment from a company, and so on. Therefore, it is desirable to position customers with a constant visit as "true loyal."

Second, the result of background intensity implies that conducting a promotion campaign at the end of the year and in early August is beneficial. In addition, setting an odd pricing to inexpensive products (especially, the products under 10 dollars) increases consumer's subconscious purchase motivation. Since we propose these strategies for promotion timing and pricing based on the background part of our model, they are effective to a wide range of customers.

Third, the results of bootstrapping indicate that it is effective to advertise to customers who get close to the first threshold (100 dollars) since the points pressure acting on the way toward the first threshold is powerful. There could be customers who are approaching the first threshold but are not influenced by the points pressure since they do not realize their proximity to it in the first place. Thus, announcing the current monthly cumulative transaction amount at the appropriate time helps ensure the proper working of the points pressure effect. Additionally, it is crucial to notify customers with the low number of monthly store visits and gold/platinum tier achievement in the past because substantiated H1 and H2 show that this kind of consumer is generally quite susceptible to the points pressure effect acting on the way toward the first threshold.

Fourth, the simulation result suggests that supermarkets should use a three-tier structure rather than a two-tier structure when implementing the customer tier program. Although there is a difference between the gold tier and the platinum tier in the major factor influencing customer's willingness to purchase (and consequently, supermarket sales), it is preferable for them to coexist in terms of the operating income.

Finally, we discuss the contribution of our study to the literature. On the one hand, as discussed in the previous sections, several attempts have been made to evaluate the effect of the customer tier program on customer purchase behavior in the context of retailing. On the other hand, there have been no previous studies that have evaluated the program while simultaneously considering purchase frequency and amount, as our study does. Our research helps us to answer questions of practical interest: "Does the customer tier program have a greater impact on either purchase frequency or purchase amount?" "What proportion of these effects ensues from changes in the points redemption rate and points pressure effects?" The results obtained from our study show that the impact on purchase frequency is larger, and the points pressure effect explains almost all of the profitability of the customer tier program for a relatively low customer tier (equivalent to "gold" in our data), while the increase in points redemption rate explains most of it for a relatively high customer tier (equivalent to "platinum" in our data). Our results suggest that, based on the characteristics of how the customer tier program affects consumer purchasing behavior, the appropriate implementation of expiration notices and points redemption rate adjustments helps to inhibit outflow to other competing stores.

2.5.2 Limitations

Our study has potential limitations. To begin with, it uses only the past purchase information as customer loyalty variables. However, demographic variables such as gender, age, income, and family composition; a qualitative survey on customer satisfaction; and information on advertising email browsing can be essential factors in defining customer loyalty. Indeed, assessment of customer loyalty and market segmentation are often conducted based on these elements across a variety of marketing domains including retailing. Here, we can give Antreas (2000), Maggioni (2016), Arminda and Sergio (2018), and Ahani et al. (2019) as appropriate examples. Since it is difficult to capture consumers' heterogeneity from their past purchase information alone, these variables allow for more flexible analysis and interpretation. Therefore, it is worthwhile to conduct further research after gathering additional information about customers.

Another limitation of our study is that it focuses only on purchase amount and does not consider any heterogeneity of products. This heterogeneity includes price, variation, inventory level, and layout, among others. The number of different products lining up in a store and the physical layout of products affect customer purchase behavior and store sales (Ton and Raman, 2010; Mowrey et al., 2018). As one example, the scan panel research enables us to consider the product characteristics. Taking the product heterogeneity into account can complicate the model and its interpretation. However, more intriguing results are likely to be obtained.

Lastly, the competitive risk structure should be considered to make our model more realistic. Consumers may change the store they usually use due to the opening of a new supermarket chain in the neighborhood, for example. Therefore, it is important to design a model considering the existence of such consumers and craft marketing strategies to improve their loyalty based on the results of this model. As a side note, the simulation result shows that our customer tier program has a stronger impact on purchase frequency than on purchase amount, suggesting that it is a disincentive for customers to visit competing supermarkets.

2.6 Appendix

2.6.1 Parameter estimation algorithm

The iterative algorithm is divided into five steps, so we itemize these procedures below. The algorithm enables us to calculate the maximum likelihood estimate $\widehat{\Theta}$. Here, we preliminarily set the values as follows: $p_{ii}^{(0)} = p_{i-}^{(0)} = 0.50$, $n_b = 20000$, $q = 1500$, $q' = 2000$, $\delta = 1/2196$.

Step 1: Set $k = 0$. Randomly choose the initial values $p_{ii}^{(0)}$ and $p_{i-}^{(0)}$ between $[0, 1]$. The former represents the probability that the event (the purchase action) i occurs naturally, and the latter signifies the stochasticity that the event i is caused by the corresponding customer covariates or the immediate events. In other words, we conveniently divide purchasing behavior into two types. Needless to say, the equation $p_{ii}^{(0)} + p_{i-}^{(0)} = 1$ holds for all $i\,(= 1, 2, \ldots, n)$.

Step 2: Based on the values of $p_{ii}^{(k)}$, randomly sample $n_b\,(< n)$ spontaneously occurring events $((t_j^b, m_j^b)_{j=1}^{n_b})$ (which we call "background event" later).

Step 3: Estimate $\nu^{(k)}$ and $\mu^{(k)}$ in the following form.

$$\nu^{(k)}(t) = \frac{1}{n} \sum_{j=1}^{n_a} \frac{1}{\sqrt{2\pi} d_j} \times \exp\left(-\frac{(t - t_j^b)^2}{2d_j^2} \right)$$
$$\mu^{(k)}(m) = \frac{1}{n} \sum_{j=1}^{n_b} \frac{1}{\sqrt{2\pi} d'_j} \times \exp\left(-\frac{(m - m_j^b)^2}{2d'^2_j} \right)$$

Here, n_a is set to one-third of n_b in this study to consider the censoring of the observation period. More specifically, if the argument t is smaller (larger) than the median of the observation period, we calculate $\nu^{(k)}$ using the n_a background events that arise after (before) the clock time t. By doing so, we can successfully estimate the background intensities near the edges of the observation period. Moreover, d_j (d'_j) denotes a Manhattan distance between t_j^b (m_j^b) and the qth (q'th) distant purchase time (amount) from t_j^b (m_j^b). The instinctive explanation for this variable bandwidth kernel density estimation is that if there are lots of background events around a certain time $\bar{t}$ that occur at a time when transaction data are accumulating, most of numerator squared terms and corresponding distances become small, resulting in high value of $\nu^{(k)}(\bar{t})$.

Step 4: Assign the estimated $\nu^{(k)}$ and $\mu^{(k)}$ to the log likelihood function. Then, maximize this function with respect to Θ using the maximum likelihood estimation. We now note that the integral calculation is handled by discretization, showing the exact form of the log likelihood function to use below.

$$l\left(\Theta\right)=\sum_{c=1}^{n_c}\sum_{i=l_c}^{l'_c}\log\lambda_c\left(t_i,m_i\mid H_{t_i}\right)-\delta\sum_{c=1}^{n_c}\sum_{u=0}^{\frac{T}{\delta}-1}\sum_{m\in\mathcal{M}}\lambda_c\left(u\delta,m\mid H_{u\delta}\right)$$

Step 5: Add 1 to k. Update $p_{ii}^{(k-1)}$ by the updating formula below and go back to Step 2. Here, we assume that event i is related to customer c.

$$\begin{aligned}p_{ii}^{(k)}&=\widehat{\xi}\nu^{(k)}\left(t_i\right)\mu^{(k)}\left(m_i\right)/\left[\widehat{\xi}\nu^{(k)}\left(t_i\right)\mu^{(k)}\left(m_i\right)+\exp\{\widehat{\alpha}_0+\widehat{\alpha}_1\left(t_i-t_{i,c-}\right)\right.\\&\quad\left.+\widehat{\alpha}_2\left(t_i-t_{i,c-}\right)^2+\widehat{\alpha}_3m_{i,c-}+\boldsymbol{x}_c^{\mathsf{T}}\widehat{\boldsymbol{\beta}}+\boldsymbol{x}_{t_i,c}^{\mathsf{T}}\widehat{\boldsymbol{\gamma}}\}\right]\end{aligned}$$

Employing this repetitive method, we can appropriately estimate a marked point process model that considers the heterogeneity based on purchase time and amount, and random effects among consumers.

2.6.2 Simulation algorithm

The thinning method is often used to simulate a focusing event in the context of point process. This is divided into four steps, so we itemize them below. The estimated parameter values are implicitly used inside the conditional intensity function appearing in the following.

Step 1: Set $\lambda^*=\max_{t\in[0,T]}\lambda_c\left(t\mid H_t\right)=\max_{t\in[0,T]}\sum_{m\in\mathcal{M}}\lambda_c\left(t,m\mid H_t\right)$. This is the maximum value over the observation period of the customer c intensity that is peripheralized with respect to purchase amount. In other words, it corresponds to the intensity at the time when the customer c is most likely to purchase.

Step 2: After initializing $i = 0$ and $t^* = 0$, perform the procedure below.

(*) Generate a random number E that follows an exponential distribution with the expectation $\frac{1}{\lambda^*}$.

Then, update $t^* \leftarrow t^* + E$.

This procedure is derived from the property of distribution followed by the inter-event interval of a stationary Poisson process.

Step 3: Set the acceptance probability $r = \lambda_c \left(t^* \mid H_{t^*}\right) / \lambda^*$ and generate an uniformly distributed random number U. If $U \leq r$, update $i \leftarrow i + 1$ and $t_i \leftarrow t^*$.

Step 4: If a candidate t^* for the purchase time is adopted in Step 3, generate a random number m^* from discrete distribution $\lambda_c \left(t^*, m \mid H_{t^*}\right) / \lambda_c \left(t^* \mid H_{t^*}\right)$ and set $m_i = m^*$. Then, repeat the steps from Step 2 (*) to Step 4 until $t^* > T$.

A series of (t_i, m_i) obtained in this way is the simulation result we wanted. Here, we note that various types of simulation are possible by controlling the influence of intervening variables inside the intensity function.

3 Empirical Research on Promotional Incentives in a Loyalty Program [4]

3.1 Introduction

The loyalty program, which is often abbreviated to LP, is a marketing strategy aiming to enhance customer loyalty and maintain good customer relationships by providing rewards for customers who frequently purchase the company's products or use its services (Dowling and Uncles, 1997). There are many academic studies on LPs, but most of them have one thing in common: they attempted to determine how an effective LP can be achieved. According to McCall and Voorhees (2010), the success of an LP is mainly determined by three factors: LP structure, reward structure, and customer fit with the LP. For example, Drèze and Nunes (2009) examined how the number and scale of tiers (a customer hierarchy determined by the store usage level) affect customers' impression of an LP with hierarchical structure. In addition, Bombaij and Dekimpe (2020) investigated how LP design, retailer strategy, and the nature of a country relate to LP success. Furthermore, Agarwal et al. (2022) surveyed the effects of perceived LP benefits and customer satisfaction with LPs on customer happiness and purchase enjoyment.

Our study investigated the short- and long-term effects of promotional incentives in an LP on customer purchase behavior. There are some previous studies on promotional incentives in an LP. For instance, Gabel and Guhl (2022) compared the effects of original LP rewards and individually targeted coupons on customer purchase frequency, purchase amount per shopping trip, and kiosk access rate by analyzing the LP of a German grocery retailer. They found that low-frequency but high-impact LP point redemptions and low-impact but high-frequency coupon redemptions are complementary in terms of LP profitability. Furthermore, Hwang et al. (2020) surveyed the relative effectiveness of two recovery options (a discount coupon for the next trip vs. bonus miles) provided to LP members in case of a service failure by an airline company. They showed that when the service failure is perceived as highly controllable by the airline company, LP members tend to consider a more concrete compensation option (a 50-dollar discount coupon in this case) as more sincere. Although previous

[4] This chapter is based on the paper "Quantifying the short- and long-term effects of promotional incentives in a loyalty program: Evidence from birthday rewards in a large retail company" written by Nishio and Hoshino in 2024. It was published in *Journal of Retailing and Consumer Services*, a peer-reviewed journal in the marketing field.

studies have focused on the differences or interaction between the effects of promotional incentives and LP rewards, there has been a notable absence of studies investigating the actual effectiveness of promotional incentives in an LP from both short- and long-term perspectives. For this reason, we performed both short- and long-term quantitative assessments of promotional incentives in an LP by applying statistical models to the LP data using birthday rewards.

From the short-term perspective, we focused on the points pressure effect. According to Taylor and Neslin (2005), this effect is defined as "the short-term impact whereby customers increase their purchase rate in an effort to earn a reward." From an empirical viewpoint, for example, Kopalle et al. (2012) revealed that hotel guests are affected by the points pressure whether the reward is a free hotel stay (price-oriented reward) or a higher customer tier enabling them to enjoy priority check-in, lounge access, and so on (service-oriented reward). In addition, from an analytical point of view, Liu et al. (2021) showed that customers' point pressure behavior can be seen in a fully rational model in which customers are assumed to be forward-looking and seek to maximize their utility function. While many LP researchers have focused on the points pressure phenomenon because elucidating the mechanism of short-term increase in customer purchase rate leads to rich managerial implications (Gandomi et al., 2019), the effects of promotional incentives in an LP on the points pressure have not been clarified yet. Therefore, our research tested a hypothesis regarding the promotional incentives and points pressure effect in an LP as described in detail in the following section.

From the long-term perspective, we focused on customer lifetime value (CLV), defined as "the estimated present value of the net benefit to the firm from the customer" (Blattberg and Deighton, 1996). CLV is widely known as a very useful metric in evaluating firms' marketing decisions in the context of customer relationship management (Gupta et al., 2006; Borle et al., 2008; Mirzaei et al., 2016). While the usefulness of CLV as a means to assess and optimize marketing strategies has been revealed in various marketing studies over the last 20 years (France and Ghose, 2019), it is pointed out that the number of LP studies that focused on the CLV is small (Chen et al., 2021). To the best of our knowledge, there is only one previous study that investigated the LP profitability from the CLV perspective. It was conducted by Gopalakrishnan et al. (2021), who examined whether the LPs without tier structure have a positive impact on CLV. The authors found that this type of LPs leads to an increase of approximately 30% in 5-year customer value, and more than 80% of the increase can be explained by the reduction in attrition rates. As considering the CLV in the LP literature is novel and managerially important, our study focused on the effects of promotional incentives in an LP on the CLV.

3.2 Hypotheses Development

We proposed the following three hypotheses, based on the points discussed in the previous section. Specifically, we considered the points pressure effect and CLV in our hypotheses formulation. The details are explained in the following subsections.

3.2.1 Points Pressure Effect

The points pressure effect is derived from the goal gradient hypothesis proposed by Hull (1932). Hull was one of the most influential psychologists of the mid-twentieth century who found in an experiment that as the rat approached the goal in the maze, its locomotion rate increased. Although this phenomenon was originally thought to occur only in the activities of primitive animals, marketing researchers in the twenty-first century demonstrated that it can be observed even in customer purchasing activities. As a pioneer, Kivetz et al. (2006) surveyed a café reward program and a website for rating songs. This group of authors indicated that the coffee purchase frequency of customers who participate in LPs is inversely proportional to the distance from the free coffee reward. In addition, they suggested that Internet users who rate songs in return for a reward become more active in their rating activities as they approach their reward goal. These human behavioral characteristics are similar to those of the rat model described above.

Recently, some studies have investigated the relationship between the strength of points pressure and reward attractiveness. For example, Nastasoiu et al. (2021) analyzed an LP implemented by an airline company, in which accumulated miles can be redeemed for 10 dollars to confirm whether the mile collection spikes around redemptions stem from the points pressure. They suggested that modest value rewards of 10 dollars are less likely to induce points pressure than more high-value ones because customers usually consider them nonchallenging and become unwilling to make more effort. In addition, Jeong et al. (2022) surveyed the relationship between rewards attached to personalized recommendation systems in the Home Meal Replacement (HMR) market and the point pressure mechanism. These authors found that an expiry date of HMR products reduced the reward value (discounts on new orders), and customers ordered them regardless of whether rewards were attached to personalized recommendations, suggesting that they were unlikely to be affected by the points pressure. Considering these previous findings, it can be said that rewards unattractive to customers lead to weak points pressure.

Following the theory proposed by Kim et al. (2021), promotional incentives in an LP are considered to lower the perceived value of original LP rewards, which implies that LP rewards become unattractive to customers after they receive promotional incentives. However, the effect of promotional incentives in an LP on the attractiveness of rewards is not permanent. So, how long does the effect last? Here, the finding from Goswami and Urminsky (2017) provides an important clue to answering this question. They indicated that customers temporarily reduce the level of their task engagement after the incentive ends, which they called "effort balancing." Applying this to the context of LP, customer motivation for attaining a reward is expected to temporarily decline after the reception of promotional incentives. To summarize the discussion so far, we can formulate the following hypothesis:

H1. After receiving promotional incentives in an LP, customers are temporarily less affected by the points pressure with the temporal decline in the attractiveness of rewards as a backdrop.

3.2.2 Customer Lifetime Value

In some previous studies on promotional incentives (not in an LP), it was pointed out that customer buying motivation is likely to be increased by them only in the short term. For example, Park et al. (2018) examined the short- and long-term effects of mobile promotions on customer purchase likelihood and expenditures. They showed that price discount coupons provided via message service on the mobile phone have a temporal but not an enduring effect that increases the customer propensity to purchase. In addition, Mutius and Huchzermeier (2021) investigated whether the design of customized category-specific coupons should be determined on the short-term performance indicators (e.g., the redemption rate) or the long-term ones (e.g., CLV). The authors found that customers whose buying behavior has accelerated in the short term owing to category-specific coupons do not necessarily remain loyal in the long term. Furthermore, Park and Yoon (2022) surveyed the effects of up-selling coupons on customers' cross-brand purchase behavior from both short- and long-term perspectives. They demonstrated that while up-selling promotions temporarily encourage customers to purchase higher-end brands, these customers tend to quickly switch back to lower-end ones.

The findings above indicate that while promotional incentives do not positively impact long-term customer purchase behavior, they do not harm perceptions of companies and can lead to immediate sales increases (Gázquez-Abad et al., 2011). This implies that such incentives can boost CLV because the temporal increase in buying motivation associated with their granting contributes to higher CLVs. However, when promotional incentives are incorporated into an LP, they temporarily lower the attractiveness of original LP rewards and weaken the points pressure, as described in the formulation of H1. As points pressure drives large profits (Taylor and Neslin, 2005; Gandomi et al., 2019), immediate sales gains from promotional incentives in an LP are likely to be exceeded by the sales decline due to weakened points pressure, resulting in lower/stable CLVs. Therefore, we can develop the following hypothesis:

H2. Promotional incentives in an LP have no positive effect of increasing the CLV.

In contrast, some preceding studies have examined the influence of promotional incentives on customers' impulsive buying behavior. Here, the impulsive buying is defined as "a result of a purchaser's immediate reaction to external stimuli that is often hedonically charged" (Madhavaram and Laverie, 2004). According to the findings from Lo et al. (2016) and Balakrishnan et al. (2020), promotional incentives work as external stimuli that induce customer impulsive actions. Specifically, the former group of authors revealed that promotional incentives with utilitarian benefits (e.g., saving money, convenience) or hedonic benefits (e.g., value, exploration) are likely to trigger the strong impulsive buying. In addition, the latter group of authors demonstrated that the increase in coupon value leads to stronger impulsiveness against the background of increased functional and psychological benefits to customers. Although the impulsive buying brings a temporal profit to a company, it was pointed out that customers tend to feel regretful after their impulsive actions (Lim et al., 2017). The regret, one of the major negative emotions customers may feel in their purchasing activity, is known to reduce their repurchase intention (Liao et al., 2017). Based on the above theoretical backgrounds, it can be inferred that customers who do not exhibit their active purchasing behavior for a certain period of time after impulsive actions induced by promotional incentives are likely to feel regretful. Considering the theory proposed by Bolton et al. (2000) according to which LP members who once regretted their action are less likely to become loyal again, the following hypothesis can be formulated:

H3. Customers who do not exhibit active purchasing behavior for a certain period of time after impulsive actions triggered by promotional incentives in an LP have particularly low CLVs.

3.3 Data

3.3.1 Data Structure and Preprocessing

To test our hypotheses, we analyzed a real-world LP using the birthday reward, which is a type of promotional incentives. We used data offered by Ryohin Keikaku Co., Ltd., a major retail company in Japan with 1136 stores as of August 2022. The data source is the same as used by Nishikawa et al. (2017). The company has domestic and foreign retail stores. Among them, we focused on domestic retail stores, often called MUJI shops. They are located in various areas, from rural to urban, and offer food, clothing, groceries, and other basic necessities. As these are private-brand products, MUJI shops have no direct competitors. We also focused on the company's online store because it has a large number of users. In May 2013, Ryohin Keikaku launched a smartphone application named MUJI Passport, which served as a membership card. This application and online store membership enabled the company to collect data about (1) customer attributes, (2) purchasing activity, and (3) application usage. Each set of data contained the following items: (1) ID number, age, sex, residential area, date of app installation, and date of registration for online store membership; (2) per-second purchase time, purchase amount, and store name; and (3) check-in history. Here, the check-in represents a customer's visit to a physical store through which they can obtain a small number of points.

We now provide a detailed explanation of our LP. Customers accumulate points called MUJI Mile (MM) to receive a reward called MUJI Shopping Point (MSP). They can obtain miles in two ways. First, 1 mile is awarded for every 1 yen (about 0.77 cents) spent, regardless of the sales channel. Second, 10 miles are provided by the check-in. Customers who accumulated a certain number of points are rewarded automatically. Specifically, 200, 300, and 500 MSPs are provided for a total of 20,000, 50,000, and 100,000 miles, respectively. Note here that 300 MSPs are granted for 30,000 miles that equal the second threshold 50,000 miles minus the first threshold 20,000 miles. The same is true for 500 MSPs. After that, 1,000 MSPs are awarded for every 100,000 miles. As MSP can be used for shopping at a rate of 1 yen per point, we can say that the point return rate of our LP is 1%. Here, it is notable that MSP is valid for only 1 month, whereas the MM continues to accumulate from the first day of March to the end day of next February. For example, a customer who obtained 300 MSPs on May 1, 2015, by reaching the second threshold of 50,000 miles, can use them in one lump or in installments to enjoy discounts on shopping expenses whenever he/she likes until June 1, 2015. In contrast, the MM balance is automatically reset to 0 at midnight on March 1 every year and keeps increasing with every purchase and check-in until the following February 28 or 29. Our LP is characterized by birthday rewards. Specifically, during the birthday month, 2 miles are awarded for every 1 yen spent. In addition, 500 MSPs are provided if customers buy a product during their birthday month. There is no product type or price restriction when earning the birthday reward of 500 MSPs. For example, even if a customer buys a small chocolate for only 70 cents during his/her birthday month, he/she can get 500 MSPs worth about 3.85 dollars. At first glance, this birthday reward seems to be too high-value a promotional incentive. However, considering the facts that the birthday reward of 500 MSPs is provided only once for each customer during his/her birthday month, and shopping usually involves costs such as travel time or shipping fees, it is considered to be of reasonable value. The birthday reward system began on March 1, 2015.

We have data from September 1, 2014 to February 29, 2016. Of these, data from March 1, 2015 to February 29, 2016, were used to test our hypotheses. Prior to the analyses, we performed three data preprocessing operations. First, we sampled customers whose age, sex, and residential areas were known. Because purchasing behavior differs depending on the values of these demographic variables, we controlled for them in our statistical models. Second, we chose customers aged 18-64 years. Third, we selected customers who had installed MUJI Passport or registered for an online store membership before September 1, 2014. The purchase data from September 1, 2014 to February 28, 2015, were used to calculate each customer's original loyalty level, so we selected customers who used MUJI shops or online store prior to September 1, 2014. Out of the original 559,637 customers, 482,728, 545,935, and 461,875 customers met conditions 1, 2, and 3, respectively. With all three conditions combined, the research sample size amounted to 394,609. Among them, 183,952 customers who returned products or canceled online purchases at any point in time between September 1, 2014, and February 29, 2016, were excluded from our analyses because it is difficult with our data to derive the exact purchase amount in each shopping of those customers. The remaining 210,657 customers either received birthday rewards in any month from March 2015 to February 2016 or did not receive them in any month. More specifically, 10,739 customers received birthday rewards in March 2015, 10,177 in April 2015, 10,722 in May 2015, 11,346 in June 2015, 10,335 in July 2015, 10,923 in August 2015, 10,694 in September 2015, 13,490 in October 2015, 11,879 in November 2015, 11,012 in December 2015, 11,013 in January 2016, 9,578 in February 2016, and 78,749 customers did not receive them in any month. To test our hypotheses, we focused only on 89,488 customers who either received birthday rewards in March 2015 or did not receive them in any month. The reason for ignoring customers who received birthday rewards in any month from April 2015 to February 2016 is that their MM balance is not always 0 when their birthday rewards are provided, which makes it difficult to capture the pure effects of promotional incentives in an LP on the points pressure and CLV.

3.3.2 Propensity Score Matching

To validate our hypotheses, we defined two customer groups; one who received birthday rewards in March 2015, and the other who did not receive them. The former is called the treatment group and the latter is the control group in the context of causal inference (Rubin, 1974; Rubin, 1978). At first glance, it seems sufficient to apply the same models to each customer group and compare their results. However, because birthday registration and purchase of any product during the birthday month are required to obtain birthday rewards, a simple comparison does not make sense because of self-selection bias.

Self-selection bias is caused by self-decisions regarding whether to participate in a study or not (Steiner et al., 2010; Kaatz, 2020). In observational studies, the subjects may not reflect the propensity of their population. For example, when conducting political surveys, many respondents are likely to be highly interested in the politics, and there is a systematic difference between the respondents and non-respondents. In our study, the treatment group is likely to have a higher original loyalty level than the control group, because loyal customers are considered highly motivated to register their birthday and purchase any product during their birthday month to obtain the birthday rewards.

Propensity score matching (PSM) is a useful tool for removing self-selection bias (Rosenbaum and Rubin, 1983). This method matches the treated and control units based on their propensity scores. After PSM, the treatment and control groups comprise similar units, and a comparison between these two groups becomes meaningful. Here, the propensity score is the conditional probability of treatment assignment given the observed covariates and is mathematically defined as follows:

$$e(\boldsymbol{x_i}) = P(z_i = 1 \mid \boldsymbol{x_i}),$$

where $\boldsymbol{x_i}$ is a vector of the observed covariates of unit i, and z_i is the treatment flag of unit i. From the definition of the propensity score, we can see that PSM is a method that brings closer the degrees of willingness of units in each group to participate in a study based on the values of their observed covariates.

PSM has often been used in the marketing literature. To cite a famous example, Wang et al. (2015) reduced the systematic differences between mobile shopping users and non-users through PSM to capture the pure effect of mobile shopping on customer purchase behavior. As observed covariates, they used past behavioral characteristics, such as the average past purchase amount and interval, and demographics, such as age, residential area, and family composition. We used as much information as possible on the observed covariates within the constraints of our data.

Table 6: Description of covariates used in PSM.

Covariate	Description
teens_flag	1 if a customer is in his/her teens; else 0
twenties_flag	1 if a customer is in his/her twenties; else 0
thirties_flag	1 if a customer is in his/her thirties; else 0
forties_flag	1 if a customer is in his/her forties; else 0
fifties_flag	1 if a customer is in his/her fifties; else 0
male_flag	1 if a customer is male; else 0
chubu_flag	1 if a customer lives in the Chubu region; else 0
chugoku_flag	1 if a customer lives in the Chugoku region; else 0
kanto_flag	1 if a customer lives in the Kanto region; else 0
kinki_flag	1 if a customer lives in the Kinki region; else 0
kyushu_flag	1 if a customer lives in the Kyushu region; else 0
shikoku_flag	1 if a customer lives in the Shikoku region; else 0
tohoku_flag	1 if a customer lives in the Tohoku region; else 0
count	the number of times a customer shopped in the pre-period
amount	customer's average purchase amount per shopping in the pre-period

Table 6 lists the observed covariates used in our matching. In this list, the items from teens_flag to tohoku_flag correspond to demographic variables (age, sex, and residential area). Here, customer age is encoded as dummy variables at 10-year intervals. In addition, the Chubu, Chugoku, Hokkaido, Kanto, Kinki, Kyushu, Shikoku, and Tohoku regions represent the area classification in Japan within which regional characteristics such as food, culture, and climate are similar. It is to be noted that sixties_flag and hokkaido_flag are excluded in advance to avoid multicollinearity. In contrast, the items count and amount correspond to each customer's original loyalty level. They are calculated using purchase data from September 1, 2014 to February 28, 2015 (pre-period) because customers' original loyalty can only be accurately estimated by their purchase behavior before treatment assignment (March, 2015).

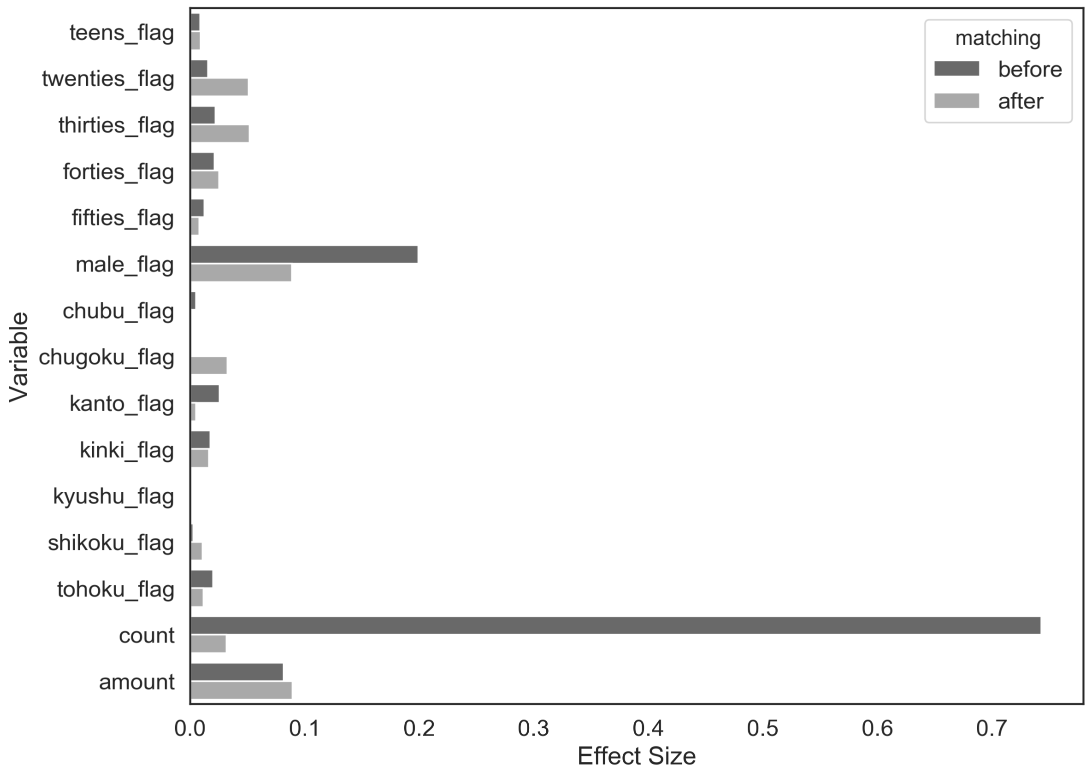

Figure 4: Cohen's d before and after PSM.

Figure 4 shows Cohen's d for each covariate before and after PSM. Here, Cohen's d is an appropriate effect size for measuring the mean difference in covariates between two groups. Specifically, this is defined as follows:

$$d_k = \frac{|\bar{x}_{k1} - \bar{x}_{k2}|}{s_k}$$
$$s_k = \sqrt{\frac{n_1 s_{k1}^2 + n_2 s_{k2}^2}{n_1 + n_2}},$$

where $\bar{x}_{kj}$ and s_{kj} are the mean and standard deviation of covariate k for group j, and n_j is the sample size of group j $(j = 1, 2)$. From this definition, we can see that Cohen's d is a standardized mean difference, and is therefore useful when the scales of each covariate are different.

According to Cohen (1988), $d_k = 0.2$ is considered a small effect size, $d_k = 0.5$ a medium effect size, and $d_k = 0.8$ a large effect size. Based on Cohen's criterion, we can assess whether our matching works well, because a valid PSM is expected to reduce the effect sizes. As shown in Figure 4, $d_{\text{male\_flag}}$ and d_{count} become particularly small after PSM. More specifically, the value of $d_{\text{male\_flag}}$ decreases from 0.199 to 0.089 and that of d_{count} decreases from 0.743 to 0.031. This implies that our matching helps to greatly reduce the differences in gender composition and original loyalty level based on purchase frequency between the treatment and control groups. By contrast, the effect sizes of the other covariates remained almost unchanged or only slightly increased in the range of 0.0 to 0.1 after PSM. Therefore, we can say that our matching is fairly effective in that the post-matching effect sizes of all covariates are less than 0.1.

Table 7: Descriptive statistics of covariates before and after PSM.

	Treatment group		Control group			
	Mean	St. Dev.	Before matching		After matching	
			Mean	St. Dev.	Mean	St. Dev.
teens_flag	0.002	0.049	0.002	0.045	0.002	0.045
twenties_flag	0.123	0.329	0.128	0.334	0.140	0.347
thirties_flag	0.495	0.500	0.484	0.500	0.469	0.499
forties_flag	0.305	0.460	0.315	0.464	0.317	0.465
fifties_flag	0.069	0.253	0.066	0.247	0.067	0.249
male_flag	0.192	0.394	0.280	0.449	0.228	0.420
chubu_flag	0.127	0.333	0.129	0.335	0.127	0.333
chugoku_flag	0.036	0.186	0.036	0.186	0.042	0.201
kanto_flag	0.514	0.500	0.527	0.499	0.517	0.500
kinki_flag	0.175	0.380	0.168	0.374	0.169	0.374
kyushu_flag	0.060	0.238	0.060	0.238	0.060	0.237
shikoku_flag	0.018	0.132	0.018	0.133	0.019	0.137
tohoku_flag	0.037	0.188	0.033	0.179	0.035	0.183
count	4.224	4.198	2.084	2.653	4.104	3.638
amount	22.533	36.616	19.214	41.240	25.901	38.413

Table 7 shows pre-matching and post-matching descriptive statistics of covariates for each group. Before PSM, the number of customers who received birthday rewards in March 2015 was 10,739, while the number of customers who did not receive them was 78,749. After implementing one-to-two matching based on propensity scores, only the latter decreased to 21,478. Therefore, the mean and standard deviation of covariates for the treatment group remain unchanged before and after PSM, as reflected in Table 7. This table shows that the pre-matching proportion of males in the control group was approximately 1.46 times greater than that in the treatment group. In addition, the average purchase frequency of customers in the control group was only approximately 49.3% of that in the treatment group before PSM. As described previously, the systematic differences between the two groups were significantly reduced using our matching method. More specifically, the post-matching proportion of males in the control group was only about 1.19 times greater than that in the treatment group. Additionally, the average purchase frequency of the control customers (4.10 times over 6 months) was almost the same as that of the treated ones (4.22 times over 6 months) after PSM. Note here that while the treated customers are guaranteed to have purchased a product in March 2015, the control ones may not have exhibited any purchasing behavior between March 1, 2015 and February 29, 2016. In fact, 2,638 out of 21,478 customers in the control group did not make a purchase in this period, and their behavior did not affect the model parameter estimation in the next section.

3.4 Model and Result

3.4.1 Points Pressure Effect

3.4.1.1 Model Structure

To strictly test our hypothesis about the points pressure effect, we applied separate models for purchase interval and purchase amount per shopping. Because these are main factors that determine retailers' profits, many preceding marketing researches focused on both of them (Morisada et al., 2019; Song et al., 2020; Li et al., 2021). As for the former, we used the Weibull proportional hazards (PH) model because the time between purchases is commonly assumed to follow a Weibull distribution (Drèze and Nunes, 2011). In this model, the density function of the interval between $(j-1)$th and jth purchases by customer i is expressed as follows:

$$p(t_{ij} \mid \alpha, \lambda_{ij}) = \alpha t_{ij}^{\alpha-1} \exp(\lambda_{ij} - \exp(\lambda_{ij}) t_{ij}^{\alpha}),$$

where α is the Weibull shape parameter and $\exp(-\lambda_{ij}/\alpha)$ is the Weibull scale parameter corresponding to the jth purchase by the customer i.

Here, by linking covariates to λ_{ij} with $\lambda_{ij} = \boldsymbol{x_{ij-1}^T \beta}$, where $\boldsymbol{x_{ij-1}}$ is a vector of covariates of the customer i at the time of the $(j-1)$th shopping trip and $\boldsymbol{\beta}$ is a vector of regression coefficients, we can obtain a hazard ratio which depends only on the values of covariates. Because the Weibull hazard function is $h(t_{ij}) = \alpha t_{ij}^{\alpha-1} \exp(\lambda_{ij})$, its ratio can be calculated as follows:

$$HR(\boldsymbol{x_{ij-1}^1}, \boldsymbol{x_{ij-1}^0}) = \frac{\alpha t_{ij}^{\alpha-1} \exp(\boldsymbol{x_{ij-1}^{1T} \beta})}{\alpha t_{ij}^{\alpha-1} \exp(\boldsymbol{x_{ij-1}^{0T} \beta})} = \exp\{\beta_k (x_{ij-1,k}^1 - x_{ij-1,k}^0)\},$$

where $\boldsymbol{x_{ij-1}^1}$ and $\boldsymbol{x_{ij-1}^0}$ are different covariate vectors for the customer i in terms of only the kth element. In other words, these vectors have the same elements except for the kth. From the above equation, we can estimate the effect of any covariate on the customer purchase interval/frequency.

In contrast, we used the generalized linear model (GLM) for the purchase amount per shopping trip because it is commonly assumed to follow a lognormal distribution (Abe, 2016). In this model, the density of the amount paid by customer i for the jth shopping is expressed as follows:

$$p(y_{ij} \mid \mu_{ij}, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2} y_{ij}} \exp\left(-\frac{(\ln y_{ij} - \mu_{ij})^2}{2\sigma^2}\right).$$

From the basic properties of the probability distributions, $y_{ij}^{'} = \ln y_{ij}$ follows a normal distribution with mean μ_{ij} and variance σ^2. Therefore, by reparameterizing μ_{ij} as $\mu_{ij} = \boldsymbol{x_{ij-1}^T \beta}$, we can place our GLM into a simple linear regression model, with $\boldsymbol{\beta}$ and σ as the unknown parameters.

Table 8 lists the covariates used in both Weibull PH model and GLM. In the list, the items from teens_flag to amount are the same as those used in our matching. Therefore, we explain the other covariates. First, to capture the effect of seasonality on purchasing behavior, dummy variables for the quarters (q1_flag, q2_flag, and q3_flag) are considered in the models. Second, to verify the hypothesis about the points pressure effect, mm_to_threshold is incorporated into the models. This covariate is defined as the number of miles remaining to earn a reward (MSPs), which enables us to capture

Table 8: Description of covariates used in Weibull PH model and GLM.

Covariate	Description
teens_flag	1 if a customer is in his/her teens; else 0
⋮	⋮
amount	customer's average purchase amount per shopping in the pre-period
q1_flag	1 for March through May; else 0
q2_flag	1 for June through August; else 0
q3_flag	1 for September through November; else 0
mm_to_threshold	remaining amount of MUJI Miles up to a threshold

changes in purchasing behavior when moving toward a goal. Based on these covariates, we separately estimated the model parameters for the two groups (the treatment and control groups). Although the simple difference-in-differences method is often used with PSM, we adopted the method of direct comparison between estimation results for each group because we wanted to apply the complex but accurate statistical models for our data.

3.4.1.2 Estimation Result

Before presenting the estimation results, we show the descriptive statistics of the dependent (purchase interval and purchase amount per shopping) and independent variables (covariates) for the treatment and control groups.

Here, we need to focus on the standard deviations (St. Dev.) of count, amount, and mm_to_threshold because we standardized these covariates prior to the estimation. Considering this, we will interpret the estimated values of the model parameters.

Table 9: Descriptive statistics of variables in Weibull PH model and GLM.

	Treatment group		Control group	
	Mean	St. Dev.	Mean	St. Dev.
purchase_interval	28.846	36.634	35.795	44.255
purchase_amount	27.052	59.553	29.369	67.661
teens_flag	0.003	0.053	0.001	0.028
twenties_flag	0.130	0.336	0.141	0.348
thirties_flag	0.506	0.500	0.502	0.500
forties_flag	0.289	0.453	0.294	0.456
fifties_flag	0.068	0.252	0.058	0.233
male_flag	0.157	0.364	0.169	0.375
chubu_flag	0.126	0.332	0.122	0.327
chugoku_flag	0.033	0.180	0.039	0.193
kanto_flag	0.532	0.499	0.544	0.498
kinki_flag	0.172	0.377	0.164	0.370
kyushu_flag	0.061	0.239	0.061	0.240
shikoku_flag	0.015	0.121	0.017	0.129
tohoku_flag	0.031	0.172	0.028	0.166
count	6.196	5.696	5.974	4.409
amount	23.911	32.161	26.337	32.449
q1_flag	0.392	0.488	0.306	0.461
q2_flag	0.225	0.418	0.224	0.417
q3_flag	0.248	0.432	0.298	0.457
mm_to_threshold	20284.773	19280.195	16372.838	14309.184

Table 10: Estimates for the Weibull PH model.

	Treatment group		Control group	
	Estimate	(St. Err.)	Estimate	(St. Err.)
Intercept	3.034^{***}	(0.067)	2.836^{***}	(0.072)
teens_flag	-0.446^{***}	(0.099)	0.153	(0.165)
twenties_flag	-0.223^{***}	(0.062)	-0.051	(0.067)
thirties_flag	-0.183^{**}	(0.062)	-0.021	(0.067)
forties_flag	-0.165^{**}	(0.062)	0.011	(0.067)
fifties_flag	-0.212^{***}	(0.063)	-0.002	(0.069)
male_flag	0.090^{***}	(0.011)	0.036^{**}	(0.011)
chubu_flag	-0.071^{**}	(0.026)	-0.046	(0.029)
chugoku_flag	-0.025	(0.033)	-0.068^{*}	(0.034)
kanto_flag	-0.117^{***}	(0.024)	-0.100^{***}	(0.027)
kinki_flag	-0.083^{**}	(0.026)	-0.068^{*}	(0.028)
kyushu_flag	-0.110^{***}	(0.029)	-0.113^{***}	(0.031)
shikoku_flag	-0.008	(0.041)	-0.089^{*}	(0.042)
tohoku_flag	0.034	(0.033)	0.049	(0.036)
count	-0.305^{***}	(0.004)	-0.217^{***}	(0.004)
amount	0.019^{***}	(0.004)	0.027^{***}	(0.004)
q1_flag	0.433^{***}	(0.014)	0.856^{***}	(0.013)
q2_flag	0.617^{***}	(0.015)	0.900^{***}	(0.014)
q3_flag	0.316^{***}	(0.014)	0.452^{***}	(0.013)
mm_to_threshold	-0.012	(0.009)	-0.044^{***}	(0.008)
male_flag : mm_to_threshold	-0.019^{+}	(0.011)	-0.026^{*}	(0.011)
q1_flag : mm_to_threshold	-0.115^{***}	(0.013)	-0.025^{+}	(0.015)
q2_flag : mm_to_threshold	-0.074^{***}	(0.013)	-0.022^{+}	(0.013)
q3_flag : mm_to_threshold	-0.058^{***}	(0.012)	-0.023^{*}	(0.010)
Weibull shape	0.774	(0.002)	0.738	(0.002)

$^{+}$ $p < .10$.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table 10 presents the estimation results of the Weibull PH model for the two groups. The number of data of the treatment group was 98330, whereas that of the control group was 104063. Because we used the LIFEREG Procedure in SAS, interpreting the estimates requires attention. Specifically, to obtain the estimate of the hazard ratio with respect to the covariate x_k, the following value must be calculated:

$$\widehat{HR}(\boldsymbol{x^1}, \boldsymbol{x^0}) = \exp\{-\hat{\alpha}\hat{\beta}_k(x_k^1 - x_k^0)\},$$

where $\widehat{\ }$ denotes the estimated value.

The following seven points can be read from Table 10. First, while there is no significant difference in purchase frequency by age group in the control group, all the estimates of the parameters corresponding to each age group from teens to 50s are significantly negative in the treatment group. Specifically, the treated customers in their teens, 20s, 30s, 40s, and 50s purchase approximately 41.23%, 18.84%, 15.22%, 13.62%, and 17.83% more frequently than those in their 60s, respectively. This tendency for younger customers to shop more frequently is observed only in the treatment group. Since all the treated customers obtained their birthday rewards that are often forgotten to receive, they tend to check e-mail notifications, including a notice on birthday rewards, more carefully compared to the control ones. In fact, Ryohin Keikaku sends to its LP members e-mail notifications regarding their registered birthday months. From the estimation result about the purchase frequency by age group, we can infer that the young generation in the treatment group is susceptible to information such as new product launches and special sales via e-mail notifications, which motivates them to shop frequently. This consideration can be supported by the theory proposed by Gilly and Zeithaml (1985), which suggests that older customers tend to have lower information-processing abilities, implying that younger customers may react to marketing communications including e-mail notifications more sensitively. Second, the purchase frequency of female customers is higher than that of male customers in both groups, with a difference of approximately 7.21% in the treatment group and 2.69% in the control group. This may be because female customers tend to have a higher preference for private-brand products offered by Ryohin Keikaku and are less likely to shop at other retail stores compared to male customers. Indeed, women are generally known as more risk-averse than men (Sundén and Surette, 1998), resulting in their tendency to purchase reliable products from familiar retail stores (Hult et al., 2019). Third, the treated (control) customers who reside in the Kanto region, the center of eastern Japan, or the Kinki region, the center of western Japan, purchase approximately 9.48% (7.66%) or 6.64% (5.15%) more frequently than the treated (control) ones who reside in the Hokkaido region, the northernmost part of Japan. This may be because MUJI shops are concentrated in metropolitan areas, making it easier for customers living in Kanto and Kinki regions to shop there. It has been pointed out that customers living in rural areas find it troublesome to go shopping for a small amount of products due to physical distances from retail stores, resulting in lower purchase frequencies (Gillespie et al., 2022). Fourth, past purchase behavior (purchase frequency and average purchase amount in the pre-period) is an important factor in explaining current purchase frequency. Specifically, when the number of times a treated (control) customer shopped in the pre-period was one more than that another treated (control) customer did, the current purchase frequency of the former is approximately 4.23% (3.70%) higher than that of the latter. Additionally, when the average purchase amount of a treated (control) customer in the pre-period was 10 dollars lower than that of another treated (control) customer, the current purchase frequency of the former is approximately 0.46% (0.62%) higher than that of the latter. Since customers' past purchase behavior is widely known to act as a predictor of their future behavior (Monsuwé et al., 2004; Weisberg et al., 2011; Li et al., 2022), our estimation result indicating that purchase frequency and average purchase amount in the pre-period have statistically significant effects on current purchase frequency is reasonable. Fifth, the purchase frequency is highest in the fourth quarter (December to February) in both groups. Specifically, customers in the treatment (control) group purchase about 39.81% (88.09%), 61.21% (94.29%), and 27.71% (39.60%) more frequently in the fourth quarter than in the first, second, and third quarters, respectively. This purchasing behavior tendency can be explained by the presence of various events, such as the Christmas party in December and New Year's first sale in January. Sixth, the estimates of the parameter corresponding to mm_to_threshold indicate that the points pressure effect about the purchase frequency is not significant in either group. In particular, the inverse points pressure effect is observed in the control group, with the purchase frequency decreasing by approximately 2.24% as the number of miles remaining up to a threshold, such as 20,000 or 50,000 miles, decreases by 10,000 miles. This inverse points pressure effect is also confirmed in the treatment group when limited to the first, second, and third quarters, with the purchase frequency decreasing by approximately 4.51%, 2.93%, and 2.30%, respectively, when the number of miles remaining up to a threshold decreases by 10,000 miles. From these values, we can see that the points pressure effect about the purchase frequency is particularly weak in the treatment group in the first quarter (March to May). Considering the fact that customers in the treatment group received their birthday rewards at some point in March, this estimation result supports H1 with respect to purchase frequency. Seventh, the sign of the parameter estimates corresponding to male_flag and mm_to_threshold interaction term suggests that in both groups, male customers are less subject to points pressure than female ones with respect to purchase frequency. This estimation result is consistent with a conventional trait of females to place more emphasis on reducing unnecessary expenses and properly managing their households compared to males (Gu et al., 2023). Since female customers are more inclined to attain rewards with monetary benefits than male ones, they may experience stronger points pressure.

Table 11: Estimates for the GLM.

	Treatment group		Control group	
	Estimate	(St. Err.)	Estimate	(St. Err.)
Intercept	7.477^{***}	(0.060)	7.379^{***}	(0.061)
teens_flag	-0.296^{***}	(0.089)	-0.422^{**}	(0.142)
twenties_flag	-0.244^{***}	(0.056)	-0.022	(0.058)
thirties_flag	-0.092^{+}	(0.055)	0.102^{+}	(0.057)
forties_flag	-0.073	(0.055)	0.125^{*}	(0.057)
fifties_flag	-0.011	(0.056)	0.142^{*}	(0.059)
male_flag	-0.005	(0.010)	0.023^{*}	(0.010)
chubu_flag	-0.031	(0.023)	-0.064^{*}	(0.025)
chugoku_flag	-0.020	(0.029)	0.043	(0.029)
kanto_flag	-0.070^{**}	(0.022)	-0.079^{***}	(0.023)
kinki_flag	-0.067^{**}	(0.023)	-0.093^{***}	(0.024)
kyushu_flag	-0.020	(0.026)	-0.054^{*}	(0.027)
shikoku_flag	0.071^{+}	(0.037)	0.017	(0.036)
tohoku_flag	0.049	(0.030)	0.046	(0.031)
count	-0.080^{***}	(0.004)	-0.078^{***}	(0.004)
amount	0.172^{***}	(0.004)	0.175^{***}	(0.004)
q1_flag	0.178^{***}	(0.012)	0.144^{***}	(0.011)
q2_flag	0.049^{***}	(0.013)	0.061^{***}	(0.012)
q3_flag	0.195^{***}	(0.013)	0.169^{***}	(0.011)
mm_to_threshold	0.111^{***}	(0.008)	0.095^{***}	(0.007)
male_flag : mm_to_threshold	-0.012	(0.010)	-0.003	(0.009)
q1_flag : mm_to_threshold	0.007	(0.011)	-0.069^{***}	(0.012)
q2_flag : mm_to_threshold	0.019^{+}	(0.011)	-0.007	(0.011)
q3_flag : mm_to_threshold	0.018^{+}	(0.010)	0.004	(0.009)
Scale	1.326	(0.006)	1.349	(0.006)

$^{+}$ $p < .10$.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 11 presents the GLM estimation result for both groups. Here, the regression coefficients represent percentages, because we applied a linear regression to the natural logarithm of the purchase amount per shopping.

The following five points can be read from Table 11. First, the treated customers in their teens to 30s have approximately 29.6%, 24.4%, and 9.2% lower average purchase amount than those in their 60s. This finding supports an earlier discussion regarding the purchasing tendencies of the younger generation in the treatment group. Second, customers in the treatment (control) group who reside in the Kanto or Kinki regions spend approximately 7.0% (7.9%) or 6.7% (9.3%) less on average than those who reside in the Hokkaido region. This is consistent with the previous discussion on the purchase tendencies of customers living in metropolitan areas. Third, past purchase behavior is an important factor in explaining current average purchase amount. Specifically, if a customer in the treatment (control) group made one more purchase in the pre-period than another customer in the same group did, the current average purchase amount of the former is approximately 1.4% (1.8%) lower than that of the latter. In addition, if a customer in the treatment (control) group spent 10 dollars less on average during the pre-period than another customer in the same group did, the current average purchase amount of the former is approximately 5.3% (5.4%) lower than that of the latter. Fourth, the average purchase amount is lowest in the fourth quarter for both groups, with customers in the treatment (control) group having approximately 17.8% (14.4%), 4.9% (6.1%), and 19.5% (16.9%) lower average purchase amounts in the fourth quarter than in the first, second, and third quarters, respectively. This finding supports the earlier discussion of purchasing tendencies in the fourth quarter. Fifth, the estimates of the parameter corresponding to mm_to_threshold indicate that the inverse points pressure effect about the purchase amount is significant in both groups. Specifically, when the number of miles remaining up to a threshold decreases by 10,000 miles, the average purchase amount of treated (control) customers decreases by approximately 5.8% (6.6%). Note that H1 does not hold with respect to purchase amount.

The estimation results suggest that birthday rewards in our LP temporarily weaken only the influence of points pressure on purchase frequency. This can be due to a reduction in motivation among LP members to make additional efforts to visit stores, either physical or online, and make purchases to earn rewards. Since going to a physical store is time-consuming and purchasing a small quantity of products from an online store involves shipping costs, the unattractive rewards following the receipt of birthday rewards can result in reduced purchase frequencies among LP members during their point-earning activities. In contrast, the points pressure effect regarding the purchase amount is not temporarily weakened by the birthday rewards probably because buying a little more than usual when visiting a physical or online store in order to earn rewards is not as hard as visiting the store itself.

3.4.2 Customer Lifetime Value

3.4.2.1 Model Structure

To compute CLV for each customer, we used the Pareto/NBD and gamma-gamma models. This approach was originally proposed by Fader et al. (2005). They suggested that in a noncontractual setting, where the time at which customers drop out is unobserved, the above models are useful in calculating customers' expected future value. Since our LP is in a noncontractual setting, these models are appropriate for estimating the CLVs of our LP members.

The Pareto/NBD model is used for predicting the purchase interval and unobserved length of customer's "lifetime." Here, the "lifetime" represents a period of time over which a customer may exhibit his/her purchasing behavior. In this model, each of the interval between purchases and lifetime length is assumed to independently follow an exponential distribution characterized by a parameter whose heterogeneity across customers is expressed in the form of gamma distribution. In contrast, the gamma-gamma model is used for predicting the average purchase amount. In this model, a series of purchase amounts are assumed to be mutually independent and identically distributed to the gamma distribution with a homogeneous shape parameter and a heterogeneous scale parameter whose heterogeneity across customers is expressed in the form of gamma distribution.

In the context of CLV, the concept of RFM is very important. Here, RFM represents recency (time of the last purchase in the observation period), frequency (number of times a customer shopped in the observation period), and monetary value (average purchase amount). Based on these terms, we can say that the Pareto/NBD model is about the recency and frequency, while the gamma-gamma model is about the monetary value.

Now, we introduce an important assumption that must be met when computing CLV using the Pareto/NBD and gamma-gamma models. It is the assumption of the independence between recency/frequency and monetary value. Under this assumption, the CLV can be divided as follows:

$$\text{CLV} = \text{margin} \times \text{revenue per transaction} \times \text{DET}.$$

The first term "margin" indicates the contribution margin ratio. It is the value of a company's sales minus its variable costs expressed as a percentage. The second term "revenue per transaction" represents an estimate of the unknown population mean of the purchase amount. This estimate can be derived from the parameter estimates in the gamma-gamma model. The third term "DET" represents the discounted expected transactions. Because CLV is the present value of the net benefit to a company from a customer, the "discount" that is proportional to the temporal distance from the present time is considered in this term. The DET can be calculated from the parameter estimates in the Pareto/NBD model. For details of their derivations, refer to Fader et al. (2005).

3.4.2.2 Model Validation

Before estimating the CLV for each customer, we checked the predictive performances of the Pareto/NBD and gamma-gamma models by dividing the observation period into two parts. The first period (estimation period) is from March 1, 2015, to October 31, 2015, and the second period (validation period) is from November 1, 2015, to February 29, 2016. As the names suggest, we estimated the model parameters using only the data during the estimation period and predicted the customer purchase behavior during the validation period.

To begin with, we checked whether the assumption of independence between recency/frequency and monetary value holds. Here, the correlation coefficient between recency and monetary value for the treatment (control) group was approximately -0.017 (0.006). In addition, the correlation coefficient between frequency and monetary value for the treatment (control) group was approximately -0.011 (-0.014). Because these values are close to 0, the independence assumption holds.

Next, we evaluated the predictive performance of the Pareto/NBD model when applied to our data. Figure 5 shows the customer averages of the observed (solid line) and predicted (dotted line) numbers of purchases made during the validation period for each group, broken down by the level of purchase frequency in the estimation period. In the Pareto/NBD model framework, the conditional expectation of the number of shopping trips during the period $(T, T+t]$ given the values of recency

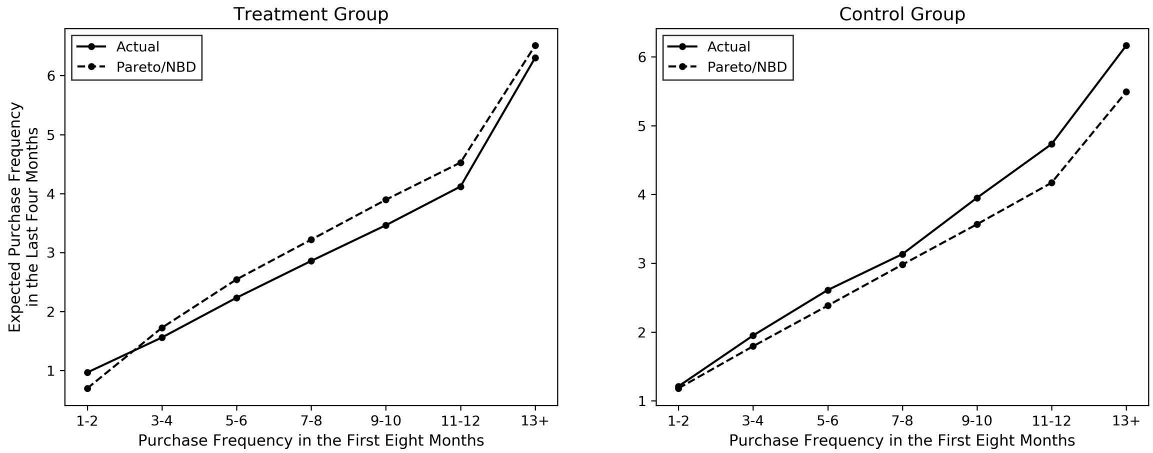

Figure 5: Observed and predicted numbers of shopping trips.

and frequency can be calculated for each customer. Since customer purchase behavior in the validation period is regarded as unknown, we can measure the predictive accuracy of the Pareto/NBD model by comparing the observed number of shopping trips during this period with the conditional expectation of the number of purchases in the same period. Figure 5 reveals that the Pareto/NBD model applied to our LP data has a considerably high predictive performance for both groups. In fact, the difference between the observed and predicted number of purchases over 4 months is only about 0.3 on average in both groups.

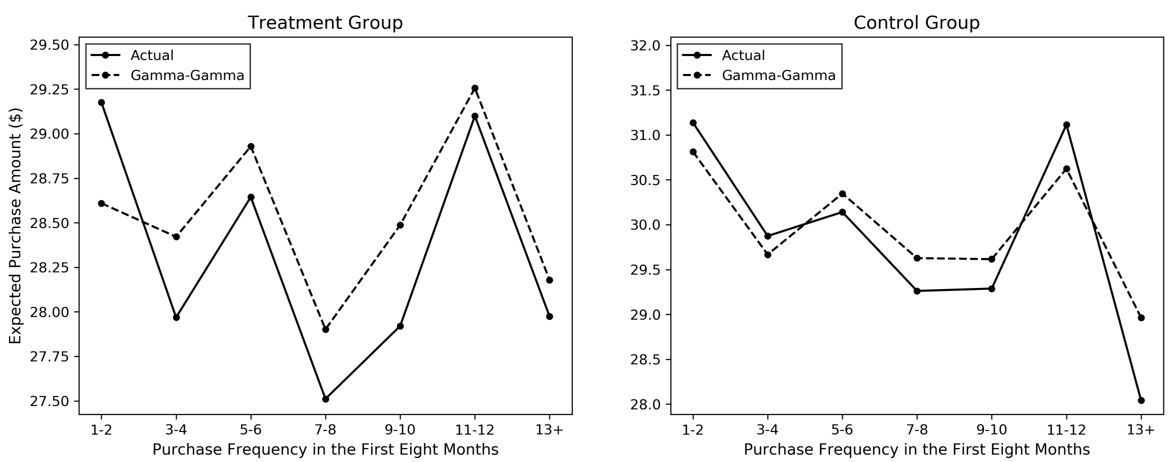

Figure 6: Observed and predicted average purchase amounts.

Finally, we evaluated the predictive performance of the gamma-gamma model when applied to our data. Figure 6 shows the customer averages of the observed (solid line) and predicted (dotted line) average purchase amounts for each group, broken down by the level of purchase frequency in the estimation period. We can now assess the predictive accuracy of the gamma-gamma model by comparing the observed average purchase amount in the observation period with the "revenue per transaction," an estimate of the population mean of the purchase amount derived from the parameter estimates in the gamma-gamma model. Figure 6 reveals that the gamma-gamma model applied to our LP data has a very high predictive performance for both groups. In fact, the difference between the observed and predicted average purchase amounts is only about 0.4 dollars on average in both groups.

3.4.2.3 Estimation Result

As we completed the validation of our models, we present and discuss the CLV estimation results. To calculate/estimate the CLV for each customer, we must set the values of "discount" and "margin" in advance. Following Fader et al. (2005) and Gupta et al. (2004), we assumed an annual discount rate of 15%. In addition, as we have no information about the actual margin of Ryohin Keikaku, we assumed a contribution margin ratio of 30%, which is consistent with the margins other marketing researchers (Reinartz and Kumar, 2000; Mutius and Huchzermeier, 2021) used.

Table 12: Median CLVs by RFM cluster.

		G=1			G=0		
M	F	R=1	R=2	R=3	R=1	R=2	R=3
1	1	20.8 (1109)	<u>77.9</u> (346)	<u>85.6</u> (195)	<u>62.1</u> (1798)	77.3 (670)	80.1 (339)
	2	70.4 (268)	<u>109.5</u> (449)	<u>117.9</u> (334)	<u>89.6</u> (529)	109.1 (762)	113.0 (589)
	3	79.6 (54)	148.2 (305)	181.0 (519)	<u>121.3</u> (139)	<u>167.5</u> (542)	<u>183.3</u> (912)
2	1	68.7 (654)	<u>115.6</u> (276)	<u>126.6</u> (143)	<u>85.1</u> (996)	113.4 (446)	117.3 (216)
	2	115.2 (270)	163.2 (477)	179.8 (346)	<u>144.0</u> (587)	<u>171.7</u> (871)	<u>181.3</u> (689)
	3	136.6 (88)	238.2 (484)	295.0 (839)	<u>204.9</u> (192)	<u>276.2</u> (867)	<u>320.4</u> (1416)
3	1	100.2 (773)	<u>183.2</u> (293)	<u>202.4</u> (164)	<u>133.9</u> (1273)	171.3 (511)	181.6 (281)
	2	193.3 (287)	281.2 (477)	295.1 (360)	<u>254.2</u> (587)	<u>316.5</u> (863)	<u>321.8</u> (631)
	3	230.1 (76)	472.7 (471)	535.4 (682)	<u>391.5</u> (179)	<u>513.5</u> (748)	<u>620.9</u> (1207)

Note: The sum of the numbers in parentheses is 10,739 (18,840) for G=1 (0).

Table 12 lists the median estimated CLVs by RFM cluster. In this table, G=1 (0) represents the treatment (control) group. The customers in each group were divided into 27 clusters based on their RFM values and the first/second tertiles of each RFM. For example, a customer in the treatment group whose recency is smaller than the first tertile of recencies in the group, frequency is between the first and second tertiles of frequencies in the group, and monetary value is larger than the second tertile of monetary values in the group is assigned to the cluster with G=1, R=1, F=2, and M=3. Here, the first and second tertiles of recency, frequency, and monetary value for the treatment (control) group are 39.380 (38.959) and 48.972 (48.096) weeks, 6 (3) and 11 (7) times, and 14.938 (15.244) and 27.860 (29.539) dollars, respectively.

The values in each cell in Table 12 represent the median estimated CLV and number of customers in the corresponding RFM cluster. After the comparison between the median CLVs in the corresponding cells in both groups, we underlined the larger one. From Table 12, it can be seen that the median CLVs of control customers exceed those of treated ones in 19 of the 27 clusters. Overall, the median CLV of all customers in the control group is approximately 166.0 dollars, while that in the treatment group is approximately 150.9 dollars, which supports H2. In addition, this table reveals that the difference between the median CLVs of treated and control customers is larger at R=1 than at R=2 or 3. Calculating the weighted average of median CLVs of treated (control) customers for each recency group, we got the results of 79.5 (123.3) dollars for R=1, 211.5 (224.5) dollars for R=2, and 274.3 (297.5) dollars for R=3. Since the ratio of the weighted average of median CLVs in the treatment group to that in the control group is particularly low at R=1, we can say that H3 may hold.

To verify H2 and H3 in more detail, we conducted three additional analyses. First, we checked the robustness with respect to H2 by changing the value of annual discount rate. Changing it from 15% to 12.5%, 17.5%, and 20%, the estimated median CLVs of all customers in the treatment (control) group became approximately 161.2 (183.9) dollars, 142.0 (151.5) dollars, and 134.3 (139.6) dollars, which strongly supports H2. Second, we recalculated the weighted averages of median CLVs of treated and control customers for R=1 after removing the cluster with R=F=M=1. The median CLV of treated customers in this cluster (20.8 dollars) is particularly low compared to that of control customers in the same cluster (62.1 dollars), and many LP members belong to this cluster (1,109 out of 10,739 for G=1 and 1,798 out of 18,840 for G=0). Hence, these median CLVs may greatly influence the weighted average of median CLVs of customers in each group for R=1. The recalculated results showed 105.9 dollars for G=1 and 147.8 dollars for G=0. As the ratio of the former to the latter (about 71.7%) is still lower than the ratios at R=2 (211.5 / 224.5 $\approx$ 94.2%) and at R=3 (274.3 /

297.5 $\approx$ 92.2%), the robustness with respect to H3 has been checked. Third, we confirmed whether impulsive actions triggered by birthday rewards in our LP lead to low CLVs. Specifically, we calculated the median CLVs of treated and control customers in the R=1 group whose total expenditures in the period from March 1, 2015, to April 15, 2015, exceed those in the period from April 16, 2015, to July 31, 2015. We could not perfectly identify impulsive actions. However, the treated customers with R=1 whose total purchase amounts in just 1.5 months including their birthday month (March) were higher than those in the following 3.5 months can be considered to have exhibited impulsive buying behavior because the increased purchasing activities around their birthdays were likely to have been induced by the birthday rewards that worked as external stimuli. The estimation results show approximately 72.2 dollars for the 1,218 treated customers and approximately 125.4 dollars for the 846 control ones. Considering the implemented one-to-two matching, the number of selected customers in the treatment group is relatively large compared to that in the control group, which supports the validity of our method for detecting impulsive buying behavior. In addition, the ratio of the former to the latter (about 57.6%) is particularly low, which strongly supports H3.

3.5 Concluding Remarks

This study investigated both short- and long-term effects of promotional incentives in an LP on customer purchase behavior by analyzing an LP using birthday rewards conducted by a major retail company in Japan. Using various statistical models (including the Weibull PH model, GLM, and Pareto/NBD & gamma-gamma models), we quantified and assessed the effects of birthday rewards, a type of promotional incentives, in terms of the points pressure effect and CLV. The analytical results revealed that promotional incentives in an LP temporarily weaken only the influence of points pressure on customer purchase frequency, have no positive effect of increasing the CLV, and induce impulsive actions among LP members that can result in particularly low CLVs. Our study was novel in that no previous research has focused on the promotional incentives in an LP that are prevalent in firms' marketing activities, giving new insights to both LP researchers and marketing managers.

3.5.1 Managerial Implications

The estimation results of our study lead to three important managerial implications. First, promotional incentives in an LP should be directly related to original LP rewards in terms of the strength of points pressure. Because the points pressure leads to short-term revenues to a company (Kivetz et al., 2006; Kopalle et al., 2012), marketing managers should design them so that customers do not become less subject to the points pressure. As described in the foundation of H1, the strength of points pressure effect and the attractiveness of original LP rewards are closely related. Therefore, promotional incentives in an LP that increase the attractiveness of LP rewards are desirable for companies in maximizing their short-term revenues. The easiest way to achieve this is to associate promotional incentives with LP rewards. For example, if LP members can receive a reward worth twice as much as what they can normally get when reaching a threshold by using a promotional incentive, the points pressure effect is likely to strengthen because the goal of attaining a reward becomes more attractive owing to the existence of a promotional incentive. Because high-value rewards are generally known to improve the LP profitability (Danaher et al., 2016), such a promotional incentive is truly effective. In addition, from the standpoint of neuroscience, high-value rewards are known to guide the high-level effort exerted by people (Massar et al., 2016; Ludwiczak et al., 2020).

Second, companies should not adopt high-value promotional incentives with little constraint in their LPs. As our analyses revealed, promotional incentives of this nature greatly diminish the attractiveness of original LP rewards, resulting in the weakened points pressure effect on the purchase frequency and lower CLVs. Therefore, promotional incentives in an LP should be either low-value with little constraint or high-value with some constraints. For example, a $1 off mobile coupon that can be used by any e-mail subscriber or a $10 off mobile coupon that is available for purchases of more than 20 dollars is desirable for companies in maintaining the high profitability of their LPs. In terms of the idiosyncratic fit heuristic proposed by Kivetz and Simonson (2003), adopting high-value promotional incentives with some constraints in an LP can help enhance the attractiveness of the LP itself. The authors revealed that under perceived idiosyncratic fit with an LP (the perception of an effort advantage over typical other customers in obtaining LP rewards), increasing the level of effort required to get the rewards is likely to enhance the perceived effort advantage and, consequently, the LP attractiveness. In addition, in terms of cognitive and affective post-purchase dissonances proposed by Balakrishnan et al. (2020), adopting low-value promotional incentives with little constraint in an LP can help keep the level of customer engagement with an LP. The authors clarified that an increase in coupon value leads to psychological or emotional discomfort for customers after any purchase action, resulting in their lowered repurchase intention. Although high-value promotional incentives with little constraint themselves can be considered not to have any negative effect on customer purchase behavior (Wierich and Zielke, 2014; Lee and Choeh, 2021), they are likely to be undesirable for companies when incorporated into their LPs.

Third, companies should design promotional incentives in an LP so that they do not induce the impulsive buying behavior. The birthday rewards in our LP (double point return rate and shopping voucher) are classified as monetary promotional incentives, which we presume to be the main reason why they triggered the impulsive actions of the LP members. Indeed, Liang and Lin (2023) demonstrated that monetary promotions (e.g., coupons, price discounts) induce impulse purchase behavior, while non-monetary ones (e.g., premiums, sweepstakes) do not affect customers' impulsiveness. In addition, Bandyopadhyay et al. (2021) indicated that monetary-immediate promotions urge customers to buy impulsively. As our analyses revealed, impulsive actions triggered by promotional incentives in an LP lead to particularly low CLVs, so monetary promotional incentives may be undesirable for retail companies. Although this managerial implication has a certain importance, our method for identifying the impulsive actions itself is not perfect due to the constraints of our data (Chen and Yao, 2018; Zhao et al., 2022). Therefore, its priority is not as high as the first and second managerial implications.

Summarizing the above managerial implications, promotional incentives that are directly related to original LP rewards, of low value with little constraint, or of high value with a certain degree of constraint are considered profitable in terms of the points pressure effect and CLV when incorporated into an LP. These multiple guidelines for designing promotional incentives in an LP based on our analytical results are useful for companies that engage in multi-purpose promotional activities.

3.5.2 Research Limitations and Future Works

This study has three potential limitations. First, we cannot perfectly identify the impulsive actions triggered by birthday rewards in our LP. Although our method for detecting impulsive buying behavior performed in Subsection 3.4.2.3 has a certain validity, a survey on customers' emotional responses to their purchasing behavior and the use of data regarding purchased products can assist verifying H3 more precisely. Many LP researches focused on the relationship between the LP effectiveness and its members' emotions (Septianto et al., 2019; Gorlier and Michel, 2020; Su et al., 2023), but there is no previous study that investigated customers' emotions such as regret and happiness in the context of promotional incentives in an LP. As described in the foundation of H3, customers tend to feel regretful after their impulsive actions. Therefore, investigating customers' emotions after a certain amount of time has elapsed since they received promotional incentives in an LP can help identify their impulse purchase behavior. In contrast, impulse buying behavior can also be determined by using the data on purchased products (Jones et al., 2003). For example, the increased purchases of hedonic products such as alcoholic drinks or cigarettes by an LP member after receiving a promotional incentive can be regarded as a result of his/her impulsive action (Xu and Huang, 2014; Deshpande et al., 2022). Thus, we put it as future works to take a survey on the LP members' emotions and obtain the data regarding purchased products.

Second, we can only use 1.5 years' worth of data for our analyses, but if we obtain data for a longer period, we could quantify the short- and long-term effects of promotional incentives in an LP distributed multiple times on customer purchase behavior. Because birthdays take place once a year, we need to conduct different analyses using at least 2.5 years' worth of data (6 months of those will be used to calculate customers' original loyalty levels) in order to capture the effects of recurrent promotional incentives in an LP. Therefore, we save it for a future work to generalize our findings in case of recurrent events.

Third, we can only use a limited number of demographic variables when implementing PSM, the Weibull PH model, and GLM. Specifically, we used data on customers' age, sex, and residential area; however, there are many other demographic variables that may influence their propensity scores and purchasing behavior. For example, job category, income, and the number of family members could be influencing factors. If we obtain data on such variables, the homogeneity of the treatment and control groups can be better ensured, and the explanatory power of the statistical models can be improved.

4 Joint Modeling of Customer Purchasing Behavior Using the Marked Hawkes Process [5]

4.1 Introduction

Consumers frequently shop at various stores in their daily lives, and their purchasing behavior can be decomposed into two key elements: purchase intervals and purchase amounts. Ideally, these two elements should be modeled simultaneously due to their mutual influence. However, in previous research, purchase intervals and amounts have often been treated separately. Additionally, there has been no prior research that formulates a model under the assumption that all past purchasing behavior of a customer can influence their current willingness to purchase. While it is natural to assume that the impact of past purchasing behavior on current willingness to purchase diminishes over time, determining how far back to consider past information when incorporating it as explanatory variables is a challenging task.

Therefore, this study applies the marked Hawkes process to ID-linked customer purchase (ID-POS) data to estimate the joint mechanism of purchase time and amount, considering all of a customer's past purchasing behavior. The marked Hawkes process has been used in various contexts, including earthquakes (Ogata, 1998), crime (Mohler et al., 2011), finance (Lee and Seo, 2017), and infectious diseases (Dong et al., 2023), to describe the relationship between event occurrence times and attribute information (marks) probabilistically and simultaneously. Unlike ARIMA and state-space models commonly used in time series analysis, a significant feature of this model is that it describes event occurrence times as random variables. In this study, the target events are each customer's purchasing behavior, and the marked Hawkes process framework is used to simultaneously estimate the mechanism related to the occurrence time (purchase time) and attribute information (purchase amount).

ID-POS data have long been used in various areas of marketing to capture customer purchasing trends and provide a basis for targeted marketing (Rossi et al., 1996). Particularly with the recent IT advancements, more detailed purchasing information, such as in-store dwell times and the contents

[5] This chapter is based on the paper "Joint modeling of customer purchasing behavior using the marked Hawkes process" written by Nishio in 2024 (in Japanese). It was published in *Japanese Journal of Applied Statistics*, a Japanese peer-reviewed journal in the field of applied statistics.

of shopping carts at specific times, has become available, further enhancing the value of ID-POS data when combined with such information (Jacobs et al., 2021). Naturally, various statistical models have been proposed for application to ID-POS data. However, regression models have been widely used in marketing literature (Wu et al., 2015; Kytö et al., 2019; Li et al., 2021; Gabel and Guhl, 2022). While the regression models excel in terms of interpretability (Halloran and Lutz, 2021), they face difficulties in simultaneously handling multiple objective variables. Additionally, when incorporating past purchasing behavior as explanatory variables, problems of arbitrariness and reduction in data quantity arise. In this respect, the marked Hawkes process, which this study focuses on, is expected to have high predictive power for heterogeneous customer purchasing behavior due to its highly flexible modeling structure (Daley and Vere-Jones, 2003). In fact, this model, which has been centrally used in the natural sciences, has never been applied in the context of retail (Reinhart, 2018; Hawkes, 2022), and this research may provide new insights for practitioners and researchers in applied statistics and marketing.

While the attempt of this study to elucidate the mechanism of customer purchasing behavior using the marked Hawkes process itself is novel, this study also introduces innovations in its model structure. To begin with, to consider customer heterogeneity (random effects), we use customer-dependent constant terms, following Halpin and Boeck (2013). Next, to include elements other than purchase intervals and amounts that may affect customer purchasing behavior, we add terms related to customer covariates, following Chiang et al. (2021). Furthermore, based on the assumption concerning purchase intervals used in Drèze and Nunes (2011), we utilize the Weibull distribution, which rarely appears in the typical framework of the marked Hawkes process. These refinements are expected to enhance the explanatory power of the marked Hawkes process employed in this study. The theoretical aspects of this model are discussed in detail in the next section.

The structure of this chapter is as follows: Section 4.2 presents the structure of the statistical model and the parameter estimation method used in this study. Section 4.3 describes the dataset to which the model is applied. Section 4.4 presents the analytical results and discusses them. Section 4.5 clarifies the contributions of this study and outlines future research directions.

4.2 Method

4.2.1 Model Structure

The marked Hawkes process is characterized by a conditional intensity function that quantifies the instantaneous likelihood of an event occurring given the history of past events (Reinhart, 2018). Therefore, we first present the conditional intensity function used in this study:

$$\lambda_c(t, m|\mathcal{H}_{ct}) = \mu_c + \sum_{t_{ci}<t, t_{ci}\in\mathcal{T}_c} R(\boldsymbol{x}_{ci}, \boldsymbol{\theta})\phi(m - m_{ci})g(t - t_{ci}).$$

This equation represents the conditional intensity function of sub-event $c \in \{1, \ldots, N\}$ when the obtained event sequence is divided into N independent subsequences (henceforth, the term "sub-" is generally omitted). For example, if the purchasing behaviors of 100 customers who shopped at a store are independent of each other, the conditional intensity function for the 50th customer's purchasing behavior is expressed as $\lambda_{50}(t, m|\mathcal{H}_{50t})$. In this study, we assume the independence of each customer's purchasing behavior, following Azeem et al. (2018).

The following provides an explanation of each symbol in the above equation. t and m represent the event occurrence time and its mark, respectively. Since the conditional intensity function is a two-variable function of t and m, it is clear that the marked Hawkes process is a method that models both time and mark simultaneously. $\mathcal{H}_{ct}$ represents the history of the occurrence of event c up to time t. Letting the occurrence time, mark, and associated covariate set of event c be $\mathcal{T}_c = \{t_{c1}, \ldots, t_{cn_c}\}, \mathcal{M}_c = \{m_{c1}, \ldots, m_{cn_c}\}, \mathcal{X}_c = \{\boldsymbol{x}_{c1}, \ldots, \boldsymbol{x}_{cn_c}\}$, respectively, we can describe $\mathcal{H}_{ct}$ concretely as follows:

$$\mathcal{H}_{ct} = \{(t_{ci}, m_{ci}, \boldsymbol{x}_{ci})|t_{ci} < t, t_{ci} \in \mathcal{T}_c, m_{ci} \in \mathcal{M}_c, \boldsymbol{x}_{ci} \in \mathcal{X}_c\}.$$

Here, n_c represents the number of times event c occurred during the observation period. Note that due to the assumption of independence between events, it is sufficient to condition only on $\mathcal{H}_{ct}$ when establishing the conditional intensity function of event c. μ_c is a constant term depending on event number c and represents the inherent likelihood of occurrence for each event. R is the term related to covariates and expresses the influence of covariates associated with event c on its occurrence. For instance, if the target events are each customer's purchasing behavior, demographic variables such as age and gender, and coupon acquisition status would correspond to the covariates. Following Chiang et al. (2021), this study defines the functional form of R as $R(\boldsymbol{x}_{ci}, \boldsymbol{\theta}) = \exp(\boldsymbol{\theta}^T \boldsymbol{x}_{ci})$. ϕ is the term related to the mark. In this study, it is defined as a power decay function:

$$\phi(m) = \begin{cases} (m + d_1)^{-q_1} & (m \geq 0) \\ (-m + d_2)^{-q_2} & (m < 0) \end{cases}.$$

This functional form has been used since Ogata (1998) and expresses the structure that events with marks deviating from past mark history are less likely to occur. For example, if a customer c typically makes purchases of around 1,000 yen, the probability of their making a 10,000 yen purchase is considered extremely low. When using the power decay function, most of $\phi(10,000 - m_{ci})$ $(i = 1, \ldots, n_c)$ will take on very small values, resulting in a small value for the corresponding conditional intensity function $\lambda_c(t, 10,000|\mathcal{H}_{ct})$. Although ϕ is typically described using the absolute value of its argument along with two parameters d and q (Reinhart, 2018), in the case of the target event in this study (customer purchasing behavior), not only the deviation width from past marks (past purchase amounts) but also their signs can serve as important elements that determine current willingness to purchase. Hence, this study introduces a conditional branching structure in ϕ as shown above. Indeed, if a customer c typically makes purchases around 1,000 yen and a customer c' around 19,000 yen, it is clear that the latter has a higher probability of making a 10,000 yen purchase. Thus, incorporating a switching structure in ϕ results in a more realistic modeling approach. g is the term related to the elapsed time since the past event occurred. In this study, it is defined as the probability density function of the Weibull distribution: $g(t) = \frac{a}{b} \left(\frac{t}{b}\right)^{a-1} \exp\left\{-\left(\frac{t}{b}\right)^a\right\}$. In previous research starting with Ogata (1988), it was often assumed that the influence from the most recent event is the strongest, and this influence attenuates over time (that is, g is a monotonically decreasing function). However, since customer purchasing behavior does not occur intensively in a short period like seismic activity, we use the probability density function of the Weibull distribution, which is widely used for modeling purchase intervals and can take on various shapes depending on the parameter values, as the functional form of g.

Given the above discussion, revisiting the conditional intensity function presented at the beginning of this subsection reveals that the instantaneous likelihood of event c with mark m at time t is quantified by adding the inherent likelihood of event c to the sum of the influences from each past event, which is determined by three elements: elapsed time, signed deviation of the mark, and related covariates. Therefore, by estimating the model parameter vector $\boldsymbol{\Theta} = (\mu_1, \ldots, \mu_N, \boldsymbol{\theta}^T, d_1, d_2, q_1, q_2, a, b)^T$, it becomes possible to simultaneously capture the mechanism related to the occurrence time and attribute information of the target events.

4.2.2 Parameter Estimation

In this study, since the marked Hawkes process uses the constant term μ_c dependent on event number c, standard maximum likelihood estimation cannot be performed. Instead, parameter estimation is conducted using the EM algorithm (Dempster et al., 1977).

First, following Daley and Vere-Jones (2003), the joint probability density function for all events in the observation period $[0, T]$ is given by

$$p_{[0,T]}(\boldsymbol{t}_{1:N}, \boldsymbol{m}_{1:N} | \boldsymbol{x}_{1:N}, \boldsymbol{\Theta}) = \prod_{c=1}^{N} \left[\prod_{i=1}^{n_c} \lambda_c(t_{ci}, m_{ci} | \mathcal{H}_{ct_{ci}}) \times \exp\left\{ - \int_0^T \lambda_c(s | \mathcal{H}_{cs}) ds \right\} \right]$$

where $\boldsymbol{t}_{1:N} = \mathcal{T}_1 \cup \cdots \cup \mathcal{T}_N, \boldsymbol{m}_{1:N} = \mathcal{M}_1 \cup \cdots \cup \mathcal{M}_N, \boldsymbol{x}_{1:N} = \mathcal{X}_1 \cup \cdots \cup \mathcal{X}_N$,

$$\lambda_c(t | \mathcal{H}_{ct}) = \sum_{m \in \overline{\mathcal{M}}} \lambda_c(t, m | \mathcal{H}_{ct}).$$

Here, $\overline{\mathcal{M}}$ represents the set of all observed marks regardless of event number. In this study, since the marks of the target events are purchase amounts, $\overline{\mathcal{M}} = \{1\text{yen}, 2\text{yen}, \ldots, m_{\max}\text{yen}\}$ is a finite set.

Based on the independence assumption between sub-events, the above joint probability density function is expressed as the product of the marginal probability density functions of each sub-event. Additionally, the marginal probability density function of event c is expressed as the product of the terms corresponding to the probability of the event occurring in specific n_c infinitesimal intervals and the terms corresponding to the probability of the event not occurring in the other infinitesimal intervals. Due to the presence of μ_c, it is impossible to find $\boldsymbol{\Theta}$ that maximizes this joint probability density function (likelihood function). Therefore, we introduce a latent variable matrix $\boldsymbol{Z}_c = (Z_{cij}) \in \{0, 1\}^{n_c \times n_c}$ concerning the internal dependency structure of event c and take the approach of maximizing the likelihood function including these variables. The (i,j) element Z_{cij} of $\boldsymbol{Z}_c$ is a binary variable that takes the value 1 only if the i-th realization of event c was caused by the $j(\leq i)$-th realization, and 0 otherwise. For example, if the 5th purchasing behavior of customer 1 was triggered by the 4th purchasing behavior, then $Z_{1,5,4} = 1$ and $Z_{1,5,j} = 0$ for $j = 1,2,3,5$. Note that since $\boldsymbol{Z}_c$ defines a one-to-one causal relationship between the realizations of event c, this matrix is a lower triangular matrix where exactly one value in each row is 1. Also, $Z_{cii} = 1$ indicates that the i-th realization of event c occurred naturally.

$$
\begin{aligned}
&p_{[0,T]}(\boldsymbol{t}_{1:N}, \boldsymbol{m}_{1:N}, \boldsymbol{z}_{1:N} | \boldsymbol{x}_{1:N}, \boldsymbol{\Theta}) \\
&= \prod_{c=1}^{N} \left[\prod_{i=1}^{n_c} \mu_c^{z_{cii}} \prod_{i=1}^{n_c} \prod_{j=1}^{i-1} \{R(\boldsymbol{x}_{cj}, \boldsymbol{\theta})\phi(m_{ci} - m_{cj}) g(t_{ci} - t_{cj})\}^{z_{cij}} \exp\left\{ -\int_0^T \lambda_c(s|\mathcal{H}_{cs}) ds \right\} \right]
\end{aligned}
$$

The above equation explicitly describes the likelihood function of the complete data, including the latent variables $\boldsymbol{Z}_{1:N} = \boldsymbol{Z}_1 \cup \cdots \cup \boldsymbol{Z}_N$. Comparing this with the likelihood function of the incomplete data shown at the beginning of this subsection, it can be seen that the conditional intensity function $\lambda_c(t_{ci}, m_{ci}|\mathcal{H}_{ct_{ci}})$ corresponding to the event c that occurred at time t_{ci} has simply changed to a product form. In fact, by incorporating the internal dependency structure of event c, it is sufficient to consider only some of the additive terms of the conditional intensity function. For instance, when $Z_{1,5,4} = 1$, since the 5th realization of event 1 is known to have been triggered by the 4th realization, only the term $R(\boldsymbol{x}_{1,4}, \boldsymbol{\theta})\phi(m_{1,5} - m_{1,4}) g(t_{1,5} - t_{1,4})$ in $\lambda_1(t_{1,5}, m_{1,5}|\mathcal{H}_{1,t_{1,5}})$ affects the likelihood of the complete data.

In the EM algorithm, parameter updating is performed in the direction that maximizes the expectation of the log-likelihood function of the complete data with respect to the posterior distribution of the latent variables ($\mathcal{Q}$ function). Therefore, referring to Veen and Schoenberg (2008), we derive the $\mathcal{Q}$ function in the setting of this study.

$$
\begin{aligned}
\mathcal{Q}(\boldsymbol{\Theta}|\boldsymbol{\Theta}^{(k)}) &= E^{\boldsymbol{Z}_{1:N}|\boldsymbol{\mathcal{H}}_{1:N},\boldsymbol{\Theta}^{(k)}}\left[\log\{p_{[0,T]}(\boldsymbol{t}_{1:N},\boldsymbol{m}_{1:N},\boldsymbol{Z}_{1:N}|\boldsymbol{x}_{1:N},\boldsymbol{\Theta})\}\middle|\boldsymbol{\mathcal{H}}_{1:N},\boldsymbol{\Theta}^{(k)}\right] \\
&= \sum_{c=1}^{N}\left[\sum_{i=1}^{n_c} p_{cii}^{(k)}\log(\mu_c) + \sum_{i=1}^{n_c}\sum_{j=1}^{i-1} p_{cij}^{(k)}\log\{R(\boldsymbol{x}_{cj},\boldsymbol{\theta})\phi(m_{ci}-m_{cj})g(t_{ci}-t_{cj})\}\right. \\
&\qquad \left. -|\overline{\mathcal{M}}|\mu_c T - \int_0^T \sum_{t_{cj}<s, t_{cj}\in\mathcal{T}_c} R(\boldsymbol{x}_{cj},\boldsymbol{\theta})\sum_{m\in\overline{\mathcal{M}}}\phi(m-m_{cj})g(s-t_{cj})ds\right] \\
&= \sum_{c=1}^{N}\left[\sum_{i=1}^{n_c} p_{cii}^{(k)}\log(\mu_c) - |\overline{\mathcal{M}}|\mu_c T + \sum_{j=1}^{n_c}\left\{\sum_{i=j+1}^{n_c} p_{cij}^{(k)}\log\{R(\boldsymbol{x}_{cj},\boldsymbol{\theta})\phi(m_{ci}-m_{cj})\right.\right. \\
&\qquad \left.\left. g(t_{ci}-t_{cj})\} - R(\boldsymbol{x}_{cj},\boldsymbol{\theta})\sum_{m\in\overline{\mathcal{M}}}\phi(m-m_{cj})\int_{t_{cj}}^{T} g(s-t_{cj})ds\right\}\right]
\end{aligned}
$$

where $\boldsymbol{\mathcal{H}}_{1:N} = \mathcal{H}_{1T}\cup\cdots\cup\mathcal{H}_{NT},\ |\overline{\mathcal{M}}|$: the number of elements in $\overline{\mathcal{M}}$, and

$$
p_{cii}^{(k)} = P(Z_{cii}=1|\boldsymbol{\mathcal{H}}_{1:N},\boldsymbol{\Theta}^{(k)}),\ p_{cij}^{(k)} = P(Z_{cij}=1|\boldsymbol{\mathcal{H}}_{1:N},\boldsymbol{\Theta}^{(k)}),\ j=1,\ldots,i-1.
$$

Additionally, in the last transformation of the equation, the order of $\sum$ and $\int$ has been changed.

Note that the EM algorithm is a method that estimates parameters by recursively repeating the E-step (a step where the above $\mathcal{Q}$ function is calculated based on the parameter estimate $\boldsymbol{\Theta}^{(k)}$ in the k-th iteration) and the M-step (a step where the parameter estimate $\boldsymbol{\Theta}^{(k+1)} = \mathrm{argmax}_{\boldsymbol{\Theta}}\ \mathcal{Q}(\boldsymbol{\Theta}|\boldsymbol{\Theta}^{(k)})$ that maximizes the $\mathcal{Q}$ function calculated in the E-step is found). According to Zhuang et al. (2002), the conditional probabilities $p_{cij}^{(k)}, p_{cii}^{(k)}$ to be calculated in the E-step are respectively

$$
p_{cij}^{(k)} = \frac{R(\boldsymbol{x}_{cj},\boldsymbol{\theta}^{(k)})\phi(m_{ci}-m_{cj};\boldsymbol{d}^{(k)},\boldsymbol{q}^{(k)})g(t_{ci}-t_{cj};a^{(k)},b^{(k)})}{\lambda_c(t_{ci},m_{ci}|\mathcal{H}_{ct_{ci}};\boldsymbol{\Theta}^{(k)})},\ p_{cii}^{(k)} = 1-\sum_{j=1}^{i-1}p_{cij}^{(k)}
$$

where $\boldsymbol{d}^{(k)} = (d_1^{(k)}, d_2^{(k)})^T,\ \boldsymbol{q}^{(k)} = (q_1^{(k)}, q_2^{(k)})^T$. The M-step then maximizes the $\mathcal{Q}$ function with these values substituted, but before describing this step, an important consideration during estimation (Schoenberg, 2013) is noted below:

> When the j-th realization of event c occurs near the end of the observation period $[0,T]$, the event pair $(i,j;c)$ is likely missing, leading to bias if used during parameter estimation.

Now, for a predetermined positive number Δ, let $J_c = \{j | t_{cj} \in [0, T-\Delta]\}$ $(c = 1, \ldots, N)$. By adding the indicator function $\mathbb{1}_{\{j \in J_c\}}$ immediately after $\sum_{j=1}^{n_c}$ in the $\mathcal{Q}$ function shown above, unbiased parameter estimation becomes possible. Here, noting that events occurring long ago have little impact, we can approximate $\int_{t_{cj}}^{T} g(s - t_{cj}) ds \approx 1$ $(j \in J_c)$, so the M-step in this study becomes as follows:

$$
\begin{aligned}
\mu_c^{(k+1)} &= \frac{\sum_{i=1}^{n_c} p_{cii}^{(k)}}{|\overline{\mathcal{M}}|T}, \; c = 1, \ldots, N \\
\boldsymbol{\theta}^{(k+1)}, \boldsymbol{d}^{(k+1)}, \boldsymbol{q}^{(k+1)} &= \mathrm{argmax}_{\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{q}} \left[\sum_{c=1}^{N} \left[\sum_{j=1}^{n_c} \mathbb{1}_{\{j \in J_c\}} \left\{ \sum_{i=j+1}^{n_c} p_{cij}^{(k)} \right. \right. \right. \\
&\quad \left. \left. \left. \log\{R(\boldsymbol{x}_{cj}, \boldsymbol{\theta}) \phi(m_{ci} - m_{cj})\} - R(\boldsymbol{x}_{cj}, \boldsymbol{\theta}) \sum_{m \in \overline{\mathcal{M}}} \phi(m - m_{cj}) \right\} \right] \right] \\
a^{(k+1)}, b^{(k+1)} &= \mathrm{argmax}_{a,b} \left[\sum_{c=1}^{N} \left[\sum_{j=1}^{n_c} \mathbb{1}_{\{j \in J_c\}} \left\{ \sum_{i=j+1}^{n_c} p_{cij}^{(k)} \log\{g(t_{ci} - t_{cj})\} \right\} \right] \right].
\end{aligned}
$$

Based on the above E-step and M-step, the parameter estimation of the marked Hawkes process in this study is performed.

4.3 Data

In this study, we demonstrate the usefulness of the marked Hawkes process described above by using ID-POS data from a supermarket provided by Trial Company, Inc. and Retail AI, Inc. Before the analysis, we extract the target customers according to the following procedure:

1. Filter customers to those in their 20s to 60s with gender information.
2. Filter customers to those who made at least one purchase per month from January to December 2021.
3. Randomly select 2,000 customers who meet the above conditions.

Step 1 is the criterion for filtering customers based on gender and age. In addition to ID-POS data, survey results related to the gender and age of some customers are also available in this study. Gender and age are important factors that determine customer purchasing behavior, so this research focuses on analyzing the working-age population with known gender information. Step 2 is the criterion for filtering customers based on purchase frequency. By extracting customers who continued to make purchases at least once a month for one year, we capture the purchasing behavior mechanism of loyal customers, which is of high interest to companies. Note that the ID-POS data from January to June 2021 (pre-period) are used for calculating customer loyalty, which will be described later, and are excluded from parameter estimation. Step 3 is intended to enhance the computational efficiency of parameter estimation. In the M-step of this study, numerical optimization is performed using the Nelder-Mead method. Since a double summation is required for each customer when calculating the objective function, $N = 2,000$ is set for computational time considerations.

Table 13: Overview of the dataframe for analysis

ID	**Time**	**Amount**	**30s**	**40s**	**50s**	**60s**	**Female**	**TotVis**	**AveAmt**	**Weekday%**
1	0.023	15.260	0	1	0	0	0	21	19.426	0.190
2	0.026	31.633	0	0	1	0	0	79	38.774	0.823
3	0.031	6.573	0	0	0	0	1	38	13.101	0.368
2	0.067	41.627	0	0	1	0	0	79	38.774	0.823
4	0.070	48.367	0	0	0	1	1	55	36.572	0.909

Table 13 provides an overview of the dataframe used for analysis. The purchase time (Time), corresponding to the event occurrence time t, is defined as the elapsed time since the start of the observation period (July 1, 2021, at 0:00:00). To prevent underflow during the calculation of the function g, the purchase time is pre-divided by 10. For example, the purchase time in the first row of Table 13 represents July 1, 2021, at 5:31:12. Similarly, the purchase amount (Amount) corresponding to the event mark m is pre-divided by 150 to prevent underflow during the calculation of the function ϕ. The three columns on the right side (TotVis, AveAmt, and Weekday%) are columns related to customer purchasing tendencies, created using ID-POS data from the pre-period. For example, for the customer with ID = 3, the number of purchases during the pre-period is 38, the average purchase amount per transaction is about 1,965 yen, and the weekday visit ratio is about 36.8%. While the original values are shown in Table 13, all three columns are standardized to facilitate the interpretation of the estimated values of the coefficient parameter vector $\boldsymbol{\theta}$ in the actual analysis. The dataframe for analysis used in this study contains 81,573 rows. Among these, purchase logs by customers in their 20s to 60s account for approximately 5.4%, 14.4%, 25.6%, 27.9%, and 26.7% of the total, respectively. Additionally, purchase logs by female customers account for about 58.8% of the total. On the other hand, the mean and standard deviation of the weekday visit ratio are approximately 0.624 and 0.199, respectively. Due to data confidentiality, summary statistics for total visits and average purchase amounts cannot be provided, so the interpretation of the corresponding coefficient parameter estimates will be based on standard deviation.

4.4 Analytical Results

4.4.1 Estimated Parameters

In this study, the EM algorithm was implemented using Python. The length of the observation period $[0, T]$ is $T = 18.4$, which is the conversion of the analysis period from July 1 to December 31, 2021, into units of 0.1 days. Additionally, considering that many customers have an average purchase interval of two weeks or more, Δ is set to 1.5.

Table 14: Parameter estimation results

	Estimate	**s.e.**
μ_{500}	0.000115	0.000003
μ_{1000}	0.000043	0.000000
μ_{1500}	0.000282	0.000004
μ_{2000}	0.000100	0.000002
θ_{30s}	0.000038	0.024
θ_{40s}	-0.010	0.016
θ_{50s}	-0.015	0.016
θ_{60s}	-0.022	0.014
θ_{female}	-0.049	0.019
θ_{count}	0.049	0.009
θ_{amount}	-0.312	0.010
θ_{weekday}	0.008	0.005
d_1	4.512	0.101
d_2	11.544	0.381
q_1	2.957	0.023
q_2	2.080	0.017
a	1.526	0.014
b	1.805	0.018

Table 14 shows the estimated values of each parameter along with their standard errors. Since there are many customer-dependent parameters μ_c, $\mu_{500}, \mu_{1000}, \mu_{1500}, \mu_{2000}$ are used as representative values in the following discussion. The coefficient parameters from $\theta_{30\text{s}}$ to θ_{weekday} correspond to the covariates: 30s flag, 40s flag, 50s flag, 60s flag, female flag, total visits during the pre-period, average purchase amount during the pre-period, and weekday visit ratio during the pre-period, respectively. Each parameter estimate represents the element of $\boldsymbol{\Theta}^{(40)}$, the estimate of $\boldsymbol{\Theta}$ at the 40th iteration of the EM algorithm. In this study, the bootstrap method (Efron, 1979) was used to calculate the standard errors associated with each parameter estimation. Specifically, the following steps were taken:

1. Perform sampling with replacement of the data $\boldsymbol{\mathcal{H}}_{1:2000}$ to obtain 2,000 samples, constructing the bootstrap sample $\tilde{\boldsymbol{\mathcal{H}}}^{(b)}_{1:2000} = \tilde{\mathcal{H}}^{(b)}_{1T} \cup \cdots \cup \tilde{\mathcal{H}}^{(b)}_{2000T}$.

2. Calculate the parameter estimates $\hat{\boldsymbol{\Theta}}^{(b)}$ under $\tilde{\boldsymbol{\mathcal{H}}}^{(b)}_{1:2000}$ based on the EM algorithm described in Subsection 4.2.2 (where the initial values and number of iterations are set to the same values).

3. Repeat the above procedure $B = 100$ times and calculate the unbiased variance of the parameter estimates obtained.

The square roots of the obtained unbiased variances were taken as the standard error estimates of each parameter. Considering execution time, the number of bootstrap iterations B was set to 100 [6], and it is known that $B = 50 - 200$ is sufficient for estimating standard errors using the bootstrap method (Kuroda, 2020).

As mentioned above, the number of iterations of the EM algorithm in this study is 40, and here, we confirm whether the algorithm has indeed converged. Figure 7 shows the transition of each parameter estimate in the EM algorithm, where the horizontal axis represents the number of iterations. To reduce the arbitrariness of the initial values, the initial values of parameters other than $\boldsymbol{d}, \boldsymbol{q}, a, b$ were all set to 0 in this study [7] ($\boldsymbol{d}, \boldsymbol{q}, a, b$'s initial values were set as $d_1^{(0)} = 4.5$, $d_2^{(0)} = 4.5$, $q_1^{(0)} = 2.5$, $q_2^{(0)} = 2.5$, $a^{(0)} = 1.2$, $b^{(0)} = 2.5$). From the general shape of the graphs shown in Figure 7, it is clear that the parameter estimates have converged to specific values at the 40th iteration. Indeed,

[6] Since the execution time for one run is about 14 hours, the entire job was divided into 10 parts and processed in parallel. The execution environment was Windows 11 Pro (OS), an 8-core 16-thread Intel Core i9-9900 (CPU), and 128GB (RAM). The total execution time was approximately 6 days.

[7] In the EM algorithm, the global convergence is not guaranteed, and parameter estimation is known to be highly dependent on the selection of initial values (Karlis and Xekalaki, 2003; Kuroda, 2020). In this study, changes in the parameter estimates due to changes in the initial values showed a tendency for only the coefficient parameters corresponding to the age group flags ($\theta_{30\text{s}}, \theta_{40\text{s}}, \theta_{50\text{s}}, \theta_{60\text{s}}$) to easily fall into local optima. However, as mentioned in the next paragraph, these parameters have large standard errors, so the problem of local optima is not considered to significantly affect the interpretation of the estimation results.

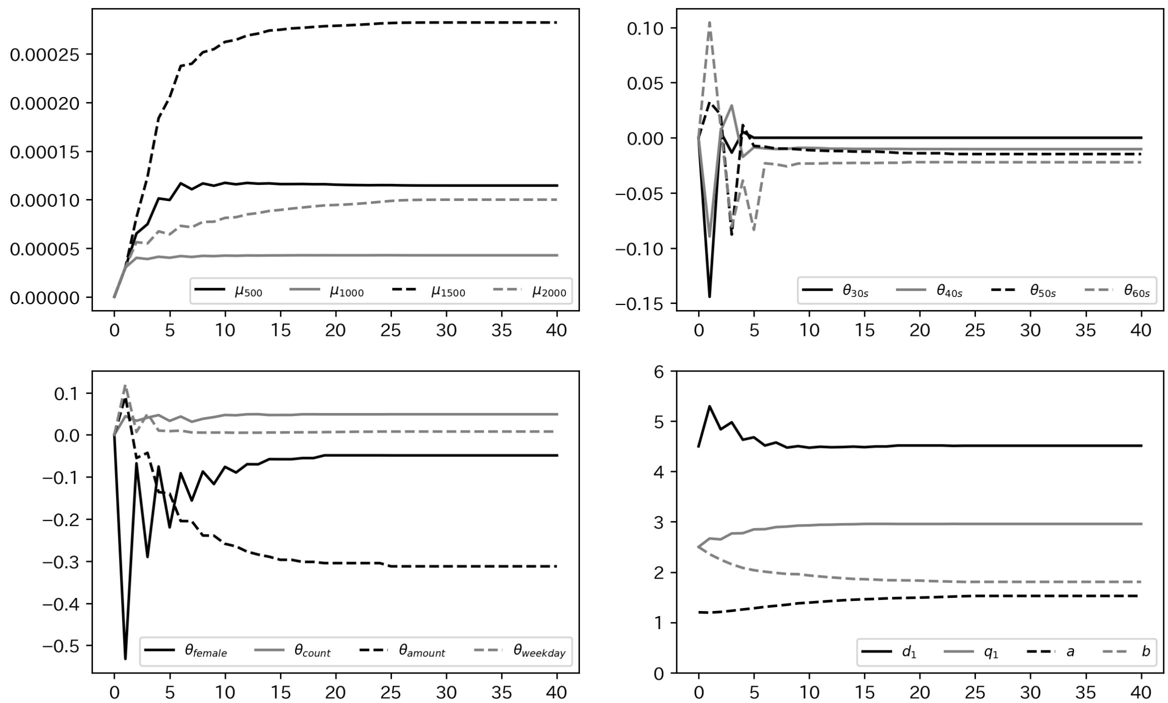

Figure 7: Transition of parameter estimates in the EM algorithm

the value $||\mathbf{\Theta}^{(40)} - \mathbf{\Theta}^{(39)}||_2^2 \approx 9.156 \times 10^{-16}$ is far below the typical convergence criterion (10^{-12}), confirming the convergence of the EM algorithm implemented in this study.

The following is a discussion of the parameter estimates shown in Table 14. First, regarding $\theta_{\text{30s}}, \theta_{\text{40s}}, \theta_{\text{50s}}, \theta_{\text{60s}}, \theta_{\text{weekday}}$, the absolute values of the estimates are less than or almost equal to their standard errors, indicating that age group and the weekday visit ratio during the pre-period do not have a significant impact on customer purchasing behavior. On the other hand, from the estimation results of $\theta_{\text{female}}, \theta_{\text{count}}, \theta_{\text{amount}}$, it is evident that male customers who have a higher total number of visits and a lower average purchase amount during the pre-period are more strongly influenced by their past purchasing behavior. For instance, a male customer who had $\bar{x}_{\text{count}}$ total visits and an average purchase amount of $\bar{x}_{\text{amount}}$ yen during the pre-period has about 25.80% higher influence from his past purchasing behavior than a female customer who had $(\bar{x}_{\text{count}} - s_{\text{count}}/2)$ total visits and an average purchase amount of $(\bar{x}_{\text{amount}} + s_{\text{amount}}/2)$ yen during the same period. Here, $\bar{x}_*$ and s_* represent the sample mean and sample standard deviation of the corresponding variable, respectively. Next, regarding $\boldsymbol{d}, \boldsymbol{q}, a, b$, the graphs in Figure 8 represent $\phi(m; \boldsymbol{d}^{(40)}, \boldsymbol{q}^{(40)})$ $(-60 \leq m \leq 60)$ and

$g(t; a^{(40)}, b^{(40)})$ $(0 < t \leq 6)$, respectively. Note that m and t are divided by 150 and 10, respectively. From the graph related to ϕ, it is evident that the probability of a customer who usually makes small purchases making a large purchase is extremely small compared to the probability of a customer who usually makes large purchases making a small purchase. For example, the influence of a past purchase with a payment of 2,000 yen on the occurrence of a 500-yen purchase event is $\phi(-10) \approx 0.0017$, whereas the influence of a past purchase with a payment of 500 yen on the occurrence of a 2,000-yen purchase event is $\phi(10) \approx 0.0004$, which is very small. On the other hand, from the graph related to g, it can be seen that the influence from past purchasing behavior rises immediately after the purchase, reaches its peak around 9.0 days later, and then gradually decays. In fact, the influence at 20 days after the purchase ($g(2) \approx 0.277$) is not significantly different from the influence at 9 days after the purchase ($g(0.9) \approx 0.415$), indicating a large variation in purchase intervals. Finally, regarding μ_c, this parameter represents the latent willingness to purchase of customer c that is not explained by the observed covariates or their past purchasing behavior. It is found that the estimates of μ_c vary significantly among customers. For example, the estimate of μ_c for the customer with ID $= 1000$ differs from that of the customer with ID $= 1500$ by approximately $2 : 13$. Although the estimates of μ_c are all small, as shown by the estimates of $\boldsymbol{\theta}$ in Table 14 and the vertical scale in Figure 8, the influence of μ_c on the conditional intensity is by no means small. As a hypothetical example, consider a female customer c' with the same estimated μ_c as the customer with ID $= 500$ (assuming that the values of the total visits and average purchase amount during the pre-period for this customer are 0 after standardization) who made only one purchase of 3,000 yen 7 days ago. Approximately 23.8% of the conditional intensity corresponding to the event of customer c' making a purchase of 3,900 yen at the present time can be explained by $\mu_{c'}^{(40)}$ [8]. The estimate of μ_c can also be interpreted as representing the latent loyalty of customer c determined by unobserved covariates such as family composition and income level, making this parameter a potentially important indicator when implementing targeted marketing in retail stores.

---

[8] The conditional intensity to be calculated is $\mu_{c'}^{(40)} + \exp(\theta_{\text{female}}^{(40)})\phi(6; \boldsymbol{d}^{(40)}, \boldsymbol{q}^{(40)})g(0.7; a^{(40)}, b^{(40)}) \approx 0.000115 + 0.952 \cdot 0.000953 \cdot 0.406 \approx 0.000483$, so the proportion occupied by $\mu_{c'}^{(40)}$ is about 23.8%.

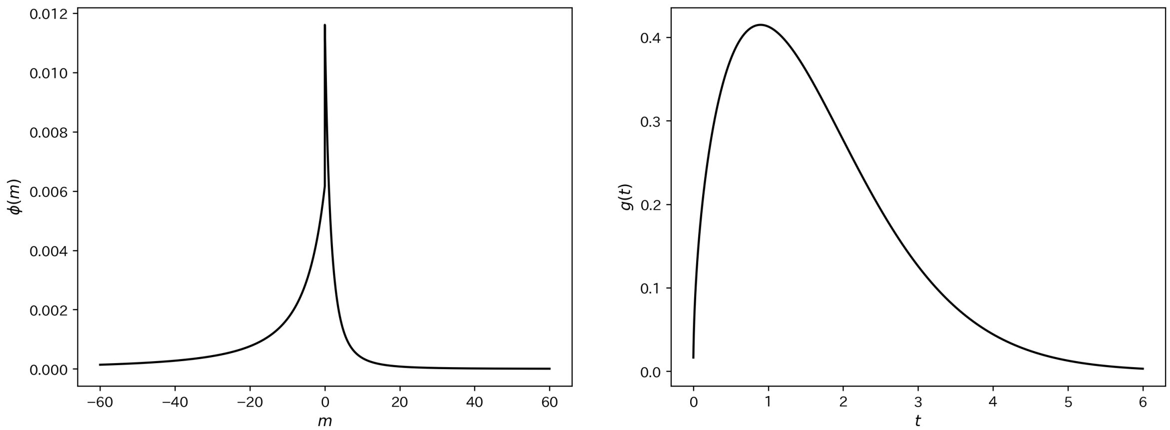

Figure 8: Plots of ϕ and g based on the estimates of $\boldsymbol{d}, \boldsymbol{q}, a, b$

4.4.2 Model Comparison

In this subsection, we compare the performance of the proposed marked Hawkes process with a traditional model to demonstrate its usefulness. Specifically, based on the estimated parameter values, simulation data are generated for the interval $(T, 1.5T]$, and the differences from actual purchase histories are measured. First, we simulate the marked Hawkes process with the conditional intensity function $\lambda_c(t, m|\mathcal{H}_{ct}; \mathbf{\Theta}^{(40)})$ $(c = 1, \ldots, N)$, using the thinning method proposed by Ogata (1981). The specific steps are as follows:

1. Initialize $i = 0,\ t_c^* = T$.
2. Update $\lambda_c^* \leftarrow \max_{t \in (t_c^*, 1.5T]} \lambda_c(t|\mathcal{H}^s_{ct}; \mathbf{\Theta}^{(40)})$ and $t_c^* \leftarrow t_c^* + \frac{1}{\lambda_c^*}$.
3. If the acceptance ratio $r = \frac{\lambda_c(t_c^*|\mathcal{H}^s_{ct_c^*}; \mathbf{\Theta}^{(40)})}{\lambda_c^*} \geq U \sim U[0, 1]$, update $i \leftarrow i + 1$ and $\tilde{t}_{ci} \leftarrow t_c^*$.
4. If an update occurs in Step 3, update $\tilde{m}_{ci} \leftarrow \mathrm{argmax}_{m \in \overline{\mathcal{M}}}\, \lambda_c(t_c^*, m|\mathcal{H}^s_{ct_c^*}; \mathbf{\Theta}^{(40)})$.
5. Repeat Steps 2 to 4 until $t_c^* > 1.5T$.

The set of purchase times $\{\tilde{t}_{c1}, \tilde{t}_{c2}, \ldots\}$ and purchase amounts $\{\tilde{m}_{c1}, \tilde{m}_{c2}, \ldots\}$ obtained by the above steps constitute the simulation data for customer c's purchasing behavior. Here, $\mathcal{H}^s_{ct}$ represents the union of the purchasing data for customer c during the analysis period $[0, T]$ and $\{(\tilde{t}_{ci}, \tilde{m}_{ci})|T < \tilde{t}_{ci} < t\}$. In addition, $\mathbf{\Theta}^{(40)}$ used in the above procedure is estimated based on ID-POS data from both the pre-period $[-T, 0)$ and the analysis period $[0, T]$. Therefore, it is important to note that no information regarding actual observations in the interval $(T, 1.5T]$ is used when generating the simulation data.

On the other hand, as a traditional model for performance comparison, this study uses a random effects model. Specifically, for the interval d_{ci} between the $(i-1)$-th and i-th purchases of customer c and the purchase amount m_{ci} of the i-th purchase by customer c, we model separately as follows:

$$d_{ci} \sim \Lambda(\beta_0 + \boldsymbol{\beta}^T \boldsymbol{x}_{c,i-1} + \gamma_c, \sigma^2),\ m_{ci} \sim \Lambda(\beta_0 + \boldsymbol{\beta}^T \boldsymbol{x}_{c,i-1} + \gamma_c, \sigma^2)$$

where $\Lambda(\cdot, \cdot)$ denotes the log-normal distribution. Here, the term corresponding to μ_c in the conditional intensity function of the marked Hawkes process proposed in this study is γ_c. To account for the most recent purchasing behavior of each customer, the covariate vector $\boldsymbol{x}_{c,i-1}$ at the $(i-1)$-th purchase time of customer c includes the rightmost eight columns of Table 13 as well as $(d_{c,i-1}, d_{c,i-2})$ and $(m_{c,i-1}, m_{c,i-2})$. Note that the traditional model does not consider purchasing behavior that occurred more than three purchases ago.

When evaluating the performance of the proposed marked Hawkes process, a commonly used model in the application scope is often chosen as a benchmark. Traditional models used for comparison typically have stronger assumptions and a limited scope of event history that can be considered within the model (Chen and Tan, 2018; Daw et al., 2024). Although point process-based models would naturally be a good choice as a traditional model, no precedent exists for applying point process-based models (including the marked Hawkes process) in the context of customer purchasing behavior (Hawkes, 2022). Therefore, among the models of customer purchasing behavior that satisfy the above characteristics, we chose the random effects model as a traditional model, known for its strong explanatory and predictive power (Gönül and Srinivasan, 1993; McKenzie et al., 2016). As mentioned in Section 4.1, regression models are primarily used in marketing literature; however, in this study, a log-linear regression model widely used in various marketing research topics is set as the traditional model (Praag and Vermeulen, 1993; Bohling et al., 2013; Chang and Zhang, 2016; Gabel and Guhl, 2022). In this model, it is assumed that purchase intervals and purchase amounts each follow a log-normal distribution independently. Additionally, the limited purchase history that can be explicitly included as explanatory variables due to data constraints makes it an appropriate model for comparison with the marked Hawkes process proposed in this study.

To estimate the parameters $(\beta_0, \beta_1, \ldots, \beta_{12}, \gamma_1, \ldots, \gamma_{2000}, \sigma^2, \sigma_\gamma^2)$ in the traditional model, this study uses the MCMC Procedure in SAS with a sampling count of 220,000, burn-in period of 20,000, and thinning interval of 5. The prior distributions for each parameter are set as follows:

$$\beta_0, \beta_1, \ldots, \beta_{12} \overset{i.i.d.}{\sim} N(0, 10^4);\ \gamma_1, \ldots, \gamma_{2000} \overset{i.i.d.}{\sim} N(0, \sigma_\gamma^2);\ \sigma^2, \sigma_\gamma^2 \overset{i.i.d.}{\sim} IG(10^{-2}, 10^{-2}).$$

Although we will not discuss the parameter estimates for the traditional model as it is beyond this study's scope, we present trace plots for selected parameters $(\beta_7, \beta_{10}, \gamma_{500}, \gamma_{1500}, \sigma^2)$ in Figure 9 (top: purchase interval model, bottom: purchase amount model) to confirm MCMC convergence. From the shape of the trace plots, it is observed that MCMC samples move within a certain range, indicating that they have sufficiently reached the stationary distribution.

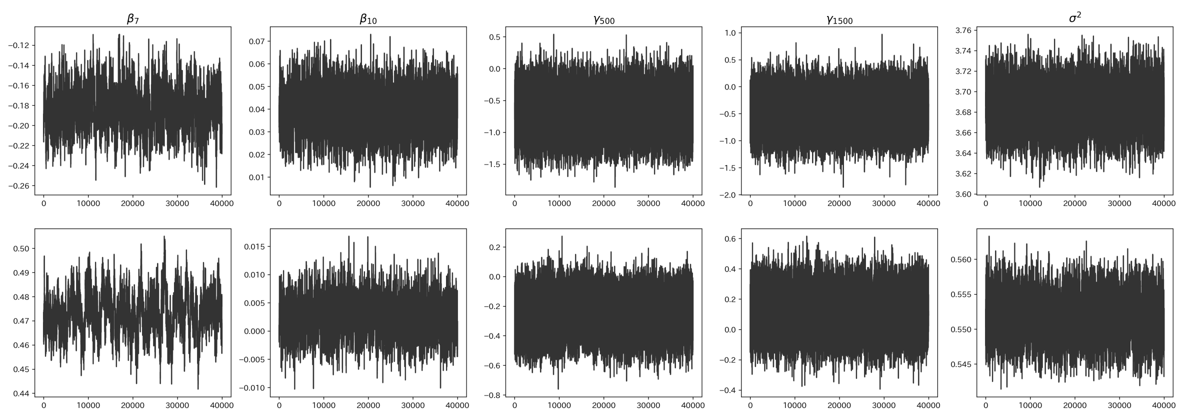

Figure 9: Trace plots in the traditional model

Letting $\bar{\theta}_{int}$ and $\bar{\theta}_{amt}$ represent the posterior mean of parameter θ for the purchase interval and purchase amount models, respectively, we can simulate the traditional model using the following steps:

1. Initialize $i = 1,\ t_c^* = t_{cn_c}$.

2. Update $\tilde{d}_{ci} \leftarrow \exp(\overline{\beta_0}_{int} + \bar{\boldsymbol{\beta}}_{int}^T \tilde{\boldsymbol{x}}_{c,i-1} + \overline{\gamma_c}_{int} + \frac{\overline{\sigma^2}_{int}}{2})$ and $t_c^* \leftarrow t_c^* + \tilde{d}_{ci}$.

3. Update $\tilde{t}_{ci} \leftarrow t_c^*$, $\tilde{m}_{ci} \leftarrow \exp(\overline{\beta_0}_{amt} + \bar{\boldsymbol{\beta}}_{amt}^T \tilde{\boldsymbol{x}}_{c,i-1} + \overline{\gamma_c}_{amt} + \frac{\overline{\sigma^2}_{amt}}{2})$, and $i \leftarrow i + 1$.

4. Repeat Steps 2 and 3 until $t_c^* > 1.5T$.

The set of purchase times $\{\tilde{t}_{c1}, \tilde{t}_{c2}, \ldots\}$ and purchase amounts $\{\tilde{m}_{c1}, \tilde{m}_{c2}, \ldots\}$ obtained here constitute the simulation data for customer c's purchasing behavior based on the traditional model. For recent purchasing behavior of customer c in $\tilde{\boldsymbol{x}}_{c,i-1}$, if it occurred before t_{cn_c} (i.e., $\leq t_{cn_c}$), the actual observed value is used, while if it occurred after t_{cn_c} (i.e., $> t_{cn_c}$), the simulated value is used. Given that parameter estimation is based solely on ID-POS data from the analysis period $[0, T]$, no information regarding actual observations in the interval $(T, 1.5T]$ is used when generating the simulation data, similar to the thinning method.

The simulation results show that for the total purchase amount M yen of 2,000 customers in the interval $(T, 1.5T]$, the predicted amount based on the traditional model shows a significant discrepancy (approximately $0.533M$ yen), while the one based on the marked Hawkes process proposed in this study achieves a high accuracy (approximately $1.016M$ yen). Given that this is a forecast of purchasing behavior over three months (January to March 2022), the high predictive power of the marked Hawkes process is evident. Major factors contributing to the large prediction error of the traditional model include: i) the modeling of purchase intervals and amounts separately, resulting in compounded prediction errors for both models, ii) the exclusion of purchasing behavior prior to three purchases ago from the explanatory variable vector for each customer, and iii) the absence of nonlinear structures represented by the power decay function ϕ and the Weibull density function g. At first glance, ii) may seem addressable by explicitly incorporating purchase behavior prior to three purchases ago as explanatory variables, but doing so would reduce the amount of usable data and may lead to missing values in the explanatory variable vector for some customers, making it difficult to fundamentally solve this issue. In this sense, the marked Hawkes process, which can consider all past purchasing behavior while simultaneously and flexibly modeling purchase time and amount, is superior to the traditional model. Its high predictive power highlights its usefulness when applied in the context of customer purchasing behavior.

4.5 Conclusion

In this study, we applied the marked Hawkes process, widely used in the natural sciences, to the context of customer purchasing in retail. This approach enabled simultaneous estimation of the mechanisms behind purchase timing and purchase amounts while considering all past purchasing behavior. Notably, two major features of this study are the estimation of parameters reflecting customer heterogeneity using the EM algorithm and the incorporation of covariates into the conditional intensity function. The simulation results demonstrated that the predictive power of the marked Hawkes process proposed here is significantly higher than the traditional modeling approach (log-linear random effects model) that treats purchase intervals and purchase amounts separately while incorporating only parts of past purchasing behavior as explanatory variables. The insights and results of this study are expected to provide new perspectives for both practitioners and applied statistics researchers.

Three main avenues for future work (research directions) can be considered. The first one is to remove the independence assumption between events. While it is reasonable to assume the independence of each customer's purchasing behavior in the supermarket context used in this study, removing this assumption would allow for modeling customer behavior in a broader range of fields. For example, in scenarios involving crowdfunding investment or airline usage, considering the correlation structure between sub-events would be more appropriate. However, without the independence assumption, the form of the joint probability density function becomes more complex, necessitating the development of more efficient parameter estimation methods. Second, it is essential to account for customer attrition within the model. This study focused on customers with continuous purchasing behavior of at least once a month, modeling both purchase times and amounts simultaneously. However, incorporating customer attrition rate into the model would likely yield more valuable practical insights. For example, capturing customer attrition could enable the estimation of customer lifetime value, an important metric in marketing. Third, incorporating competitive risk structure into the model is important. Since competing stores exist for most retailers, analysts can observe only a subset of all customer purchasing behavior. By explicitly integrating the competitive risk structure into the conditional intensity function, it would be possible to achieve a more realistic modeling of customer purchasing behavior.

5 Conclusion of the Dissertation

This dissertation presents new methodologies in the fields of marketing science and econometric marketing. Specifically, the first and third studies focused on the marked point process approaches which have not been applied to customer purchase behavior, and the second study quantified the long-term effect of promotional incentives in a loyalty program on customer lifetime value using the Pareto/NBD and gamma-gamma models which have not been used in prior studies on loyalty programs. The three empirical studies revealed that there are various econometric models that have not been extensively studied, but have high explanatory and predictive powers when applied to ID-linked customer purchase data in retail stores.

The first study (Chapter 2) showed that initially loyal customers are less susceptible to the points pressure, the points pressure towards the first threshold is stronger than towards the second threshold in the three-tier customer tier program, and the three-tier reward program is superior to the two-tier one in terms of operating income for retail stores. The second study (Chapter 3) demonstrated that customers become temporarily less affected by the points pressure after obtaining promotional incentives in a loyalty program, such incentives have no positive effect on customer lifetime value, and impulsive buying triggered by these incentives can lead to particularly low customer lifetime values. These empirical studies contribute to the literature on loyalty programs by quantitatively revealing the profitability of customer tier programs and promotional incentives within a loyalty program. On the other hand, the third study (Chapter 4) indicated that the marked Hawkes process greatly outperforms the log-linear regression model with the random intercept for long-term forecasting of customer purchasing behavior. This empirical study contributes to the literature on consumer behavior and point processes by extending the range of application of point process approaches to consumer purchasing behavior.

Customer purchase behavior is difficult to understand with a background of heterogenous purchasing tendencies, implementation of various marketing strategies by firms including competing ones, and incidental events such as bad weather or urgent business. Therefore, the incorporation of random effects terms, within-customer correlated variance terms, and observable variables related to marketing strategies and incidental events into the econometric models used is essential for accurate estimation and prediction of consumer purchasing behavior. In this regard, the empirical studies in this dissertation still offer opportunities for further refinement. As a direction for future research, the use of hierarchical bayesian models and non-parametric marked point process models could make these studies both more practical and academically rigorous. For example, using the hierarchical model with demographic variables such as gender, age, and residential area would enable the estimation of the strength of the points pressure effect for each customer group with different attributes, resulting in more detailed practical implications. In addition, not restricting the marked point process model to a specific parametric structure can lead to more flexible and accurate understanding of customer purchasing behavior because heterogenous purchasing tendencies are difficult to capture in an arbitrary parametric form. For the developments in the fields of marketing science and econometric marketing, it is necessary to further investigate customer purchase behavior under both academically and practically meaningful research questions through the application of various sophisticated econometric models.

Acknowledgments

First, I would like to sincerely thank my supervisor, Professor Takahiro Hoshino. His enthusiastic research guidance based on his extensive knowledge of econometrics, marketing science, behavioral economics, and psychology has been indispensable to my research activities at the Graduate School of Economics, Keio University. Also, I thank my fellow graduate students in Hoshino Seminar for the discussions on marketing science and econometric marketing.

I thank Japan Science and Technology Agency (JST) and Japan Society for the Promotion of Science (JSPS) for financially supporting me during my doctoral studies. I would also like to express my gratitude to Trial Company, Inc., Retail AI, Inc., and Ryohin Keikaku Co., Ltd. for providing me with valuable ID-linked customer purchase data.

Finally, I deeply thank my parents, Naoyoshi Nishio and Yoshiko Nishio, and my grandparents for their love and unwavering support.

References

Abe, M., 2016. Deriving Customer Lifetime Value from RFM Measures: Insights into Customer Retention and Acquisition. From Little's Law to Marketing Science: Essays in Honor of John D.C. Little. 127-169. https://mitpress.mit.edu/9780262029919/from-littles-law-to-marketing-science/.

Agarwal, R., Mehrotra, A., Misra, D., 2022. Customer happiness as a function of perceived loyalty program benefits - A quantile regression approach. Journal of Retailing and Consumer Services. 64, 102770. https://doi.org/10.1016/j.jretconser.2021.102770.

Ahani, A., Nilashi, M., Ibrahim, O., Sanzogni, L., Weaven, S., 2019. Market segmentation and travel choice prediction in spa hotels through TripAdvisor's online reviews. International Journal of Hospitality Management. 80, 52-77. https://doi.org/10.1016/j.ijhm.2019.01.003.

Antreas, D.A., 2000. Customer satisfaction cues to support market segmentation and explain switching behavior. Journal of Business Research. 47, 191-207. https://doi.org/10.1016/S0148-2963(98)00060-5.

Arminda, A.S., Sergio, M.G., 2018. Understanding tourism loyalty: horizontal vs. destination loyalty. Tourism Management. 65, 245-255. https://doi.org/10.1016/j.tourman.2017.10.011.

Azeem, M.M., Baker, D., Villano, R.A., Mounter, S., Griffith, G., 2018. Food shoppers' share of wallet: A small city case in a changing competitive environment. Journal of Retailing and Consumer Services. 43, 119-130. https://doi.org/10.1016/j.jretconser.2018.03.009.

Balakrishnan, J., Foroudi, P., Dwivedi, Y.K., 2020. Does online retail coupons and memberships create favourable psychological disposition? Journal of Business Research. 116, 229-244. https://doi.org/10.1016/j.jbusres.2020.05.039.

Bandyopadhyay, N., Sivakumaran, B., Patro, S., Kumar, R.S., 2021. Immediate or delayed! Whether various types of consumer sales promotions drive impulse buying?: An empirical investigation. Journal of Retailing and Consumer Services. 61, 102532. https://doi.org/10.1016/j.jretconser.2021.102532.

Bazargana, A., Karray, S., Zolfagharia, S., 2017. Modeling reward expiry for loyalty programs in a competitive market. International Journal of Production Economics. 193, 352-364. https://doi.org/10.1016/j.ijpe.2017.08.001.

Bijwaard, G.E., Franses, P.H., Paap, R., 2006. Modeling purchases as repeated events. Journal of

Business & Economic Statistics. 24, 487-502. https://doi.org/10.1198/073500106000000242.

Blattberg, R.C., Deighton, J., 1996. Manage marketing by the customer equity test. Harvard Business Review. 74, 136-144.

Blattberg, R.C., Kim, B., Neslin, S.A., 2008. Database Marketing: Analyzing and Managing Customers. Springer, New York.

Bohling, T.R., Kumar, V., Shah, R., 2013. Predicting purchase timing, product choice, and purchase amount for a firm's adoption of a radically innovative information technology: An analysis of cloud computing services. Service Science. 5, 102-123. https://doi.org/10.1287/serv.1120.0039.

Bolton, R.N., Kannan, P.K., Bramlett, M.D., 2000. Implications of loyalty program membership and service experiences for customer retention and value. Journal of the Academy of Marketing Science. 28, 95-108. https://doi.org/10.1177/0092070300281009.

Bombaij, N.J.F., Dekimpe, M.G., 2020. When do loyalty programs work? The moderating role of design, retailer-strategy, and country characteristics. International Journal of Research in Marketing. 37, 175-195. https://doi.org/10.1016/j.ijresmar.2019.07.003.

Borle, S., Singh, S.S., Jain, D.C., 2008. Customer lifetime value measurement. Management Science. 54, 100-112. https://doi.org/10.1287/mnsc.1070.0746.

Brantingham, P.J., Yuan, B., Herz, D., 2021. Is gang violent crime more contagious than non-gang violent crime? Journal of Quantitative Criminology. 37, 953-977. https://doi.org/10.1007/s10940-020-09479-1.

Bridson, K., Evans, J., Hickman, M., 2008. Assessing the relationship between loyalty program attributes, store satisfaction and store loyalty. Journal of Retailing and Consumer Services. 15, 364-374. https://doi.org/10.1016/j.jretconser.2007.08.004.

Chang, C.-W., Zhang, J.Z., 2016. The effects of channel experiences and direct marketing on customer retention in multichannel settings. Journal of Interactive Marketing. 36, 77-90. https://doi.org/10.1016/j.intmar.2016.05.002.

Cheema, A., Bagchi, R., 2011. The effect of goal visualization on goal pursuit: implications for consumers and managers. Journal of Marketing. 75, 109-123. https://doi.org/10.1509/jmkg.75.2.109.

Chen, C.C., Yao, J.Y., 2018. What drives impulse buying behaviors in a mobile auction? The perspective of the Stimulus-Organism-Response model. Telematics and Informatics. 35, 1249-1262.

https://doi.org/10.1016/j.tele.2018.02.007.

Chen, F., Tan, W.H., 2018. Marked self-exciting point process modelling of information diffusion on Twitter. Annals of Applied Statistics. 12, 2175-2196. https://doi.org/10.1214/18-AOAS1148.

Chen, Y., Mandler, T., Meyer-Waarden, L., 2021. Three decades of research on loyalty programs: A literature review and future research agenda. Journal of Business Research. 124, 179-197. https://doi.org/10.1016/j.jbusres.2020.11.057.

Chhabra, S., 2017. An empirical analysis of the effect of a retailers loyalty programme on their customers' loyalty. Global Business Review. 18, 445-464. https://doi.org/10.1177/0972150916668612.

Chiang, W.-H., Liu, X., Mohler, G., 2021. Hawkes process modeling of COVID-19 with mobility leading indicators and spatial covariates. International Journal of Forecasting. 38, 505-520. https://doi.org/10.1016/j.ijforecast.2021.07.001.

Christensen, T.M., Hurn, A.S., Lindsay, A.J., 2012. Forecasting spikes in electricity prices. International Journal of Forecasting. 28, 400-411. https://doi.org/10.1016/j.ijforecast.2011.02.019.

Clements, A.E., Herrera, R., Hurn, A.S., 2015. Modelling interregional links in electricity price spikes. Energy Economics. 51, 383-393. https://doi.org/10.1016/j.eneco.2015.07.014.

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences (2nd ed.). https://doi.org/10.4324/9780203771587.

Daley, D.J., Vere-Jones, D., 2003. An Introduction to the Theory of Point Processes, Volume I: Elementary Theory and Methods. Springer, New York.

Danaher, P.J., Sajtos, L., Danaher, T.S., 2016. Does the reward match the effort for loyalty program members? Journal of Retailing and Consumer Services. 32, 23-31. http://dx.doi.org/10.1016/j.jretconser.2016.05.015.

Daw, A., Castellanos, A., Yom-Tov, G.B., Pender, J., Gruendlinger, L., 2024. The co-production of service: Modeling services in contact centers using Hawkes processes. Management Science. https://doi.org/10.1287/mnsc.2021.04060.

Demoulin, V., Davison, A.C., McNeil, A.J., 2005. Estimating value-at-risk: a point process approach. Quantitative Finance. 5, 227-234. https://doi.org/10.1080/14697680500039613.

Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society: Series B (Methodological). 39, 1-22.

https://doi.org/10.1111/j.2517-6161.1977.tb01600.x.

Deshpande, B., Pradhan, D., Sivakumaran, B., Lyngdoh, T., 2022. The impact of advertising appeals on impulse buying. Marketing Intelligence and Planning. 40, 358-371. https://doi.org/10.1108/MIP-11-2021-0390.

Dong, Z., Zhu, S., Xie, Y., Mateu, J., Rodríguez-Cortés, F.J., 2023. Non-stationary spatio-temporal point process modeling for high-resolution COVID-19 data. Journal of the Royal Statistical Society Series C: Applied Statistics. 72, 368-386. https://doi.org/10.1093/jrsssc/qlad013.

Dorotic, M., Verhoefab, P.C., Fok, D., Bijmoltb, T.H., 2014. Reward redemption effects in a loyalty program when customers choose how much and when to redeem. International Journal of Research in Marketing. 31, 339-355. https://doi.org/10.1016/j.ijresmar.2014.06.001.

Dowling, G.R., Uncles, M., 1997. Do customer loyalty programs really work? Sloan Management Review. 38.

Drèze, X., Nunes, J.C., 2009. Feeling superior: The impact of loyalty program structure on consumers' perceptions of status. Journal of Consumer Research. 35, 890-905. https://doi.org/10.1086/593946.

Drèze, X., Nunes, J.C., 2011. Recurring goals and learning: The impact of successful reward attainment on purchase behavior. Journal of Marketing Research. 48, 268-281. https://doi.org/10.1509/jmkr.48.2.268.

Efron, B., 1979. Bootstrap methods: Another look at the Jackknife. Annals of Statistics. 7, 1-26. https://doi.org/10.1214/aos/1176344552.

Embrechts, P., Liniger, T., Lin, L., 2011. Multivariate Hawkes processes: an application to financial data. Journal of Applied Probability. 48, 367-378. https://doi.org/10.1239/jap/1318940477.

Fader, P.S., Hardie, B.G.S., Lee, K.L., 2005. RFM and CLV: Using iso-value curves for customer base analysis. Journal of Marketing Research. 42, 415-430. https://doi.org/10.1509/jmkr.2005.42.4.415.

Fraccaro, A., Macé, S., Parguel, B., 2021. The not-so-odd couple: odd pricing in a luxury context. Journal of Business Research. 136, 356-365. https://doi.org/10.1016/j.jbusres.2021.07.048.

France, S.L., Ghose, S., Marketing analytics: Methods, practice, implementation, and links to other fields. Expert Systems with Applications. 119, 456-475. https://doi.org/10.1016/j.eswa.

2018.11.002.

Gabel, S., Guhl, D., 2022. Comparing the effectiveness of rewards and individually targeted coupons in loyalty programs. Journal of Retailing. 98, 395-411. https://doi.org/10.1016/j.jretai.2021.08.001.

Gandomi, A., Zolfaghari, S., 2013. Profitability of loyalty reward programs: an analytical investigation. Omega. 41, 797-807. https://doi.org/10.1016/j.omega.2012.10.003.

Gandomi, A., Bazargan, A., Zolfaghari, S., 2019. Designing competitive loyalty programs: a stochastic game-theoretic model to guide the choice of reward structure. Annals of Operations Research. 280, 267-298. https://doi.org/10.1007/s10479-019-03179-1.

Gázquez-Abad, J.C., Canniére, M.H.D., Martínez-López, F.J., 2011. Dynamics of customer response to promotional and relational direct mailings from an apparel retailer: The moderating role of relationship strength. Journal of Retailing. 87, 166-181. https://doi.org/10.1016/j.jretai.2011.03.001.

Gillespie, R., DeWitt, E., Trude, A.C.B., Haynes-Maslow, L., Hudson, T., Anderson-Steeves, E., Barr, M., Gustafson, A., 2022. Barriers and facilitators of online grocery services: Perceptions from rural and urban grocery store managers. Nutrients. 14, 3794. https://doi.org/10.3390/nu14183794.

Gilly, M.C., Zeithaml, V.A., 1985. The elderly consumer and adoption of technologies. Journal of Consumer Research. 12, 353-357. https://doi.org/10.1086/208521.

Gönül, F., Srinivasan, K., 1993. Consumer purchase behavior in a frequently bought product category: Estimation issues and managerial insights from a hazard function model with heterogeneity. Journal of the American Statistical Association. 88, 1219-1227. https://doi.org/10.1080/01621459.1993.10476401.

Gopalakrishnan, A., Jiang, Z., Nevskaya, Y., Thomadsen, R., 2021. Can non-tiered customer loyalty programs be profitable? Marketing Science. 40, 508-526. https://doi.org/10.1287/mksc.2020.1268.

Gorlier, T., Michel, G., 2020. How special rewards in loyalty programs enrich consumer-brand relationships: The role of self-expansion. Psychology and Marketing. 37, 588-603. https://doi.org/10.1002/mar.21328.

Goswami, I., Urminsky, O., 2017. The dynamic effect of incentives on postreward task engagement. Journal of Experimental Psychology: General. 146, 1-19. http://dx.doi.org/10.1037/

xge0000206.

Gu, C., Hu, L., Lei, X., Yang, D., 2023. Do electronic coupon-using behaviors make men womanish? The effect of the coupon-feminine stereotype. Journal of Theoretical and Applied Electronic Commerce Research. 18, 1637-1659. https://doi.org/10.3390/jtaer18030083.

Gupta, S., Lehmann, D.R., Stuart, J.A., 2004. Valuing customers. Journal of Marketing Research. 41, 7-18. https://doi.org/10.1509/jmkr.41.1.7.25084.

Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N., Ravishanker, N., Sriram, S., 2006. Modeling customer lifetime value. Journal of Service Research. 9, 139-155. https://doi.org/10.1177/1094670506293810.

Halloran, T.J., Lutz, R.J., 2021. Let's give them something to talk about: Which social media engagements predict purchase frequency? Journal of Interactive Marketing. 56, 83-95. https://doi.org/10.1016/j.intmar.2021.05.003.

Halpin, P.F., Boeck, P.D., 2013. Modelling dyadic interaction with Hawkes processes. Psychometrika. 78, 793-814. https://doi.org/10.1007/s11336-013-9329-1.

Hawkes, A.G., 2022. Hawkes jump-diffusions and finance: A brief history and review. The European Journal of Finance. 28, 627-641. https://doi.org/10.1080/1351847X.2020.1755712.

Hendel, I., Nevo, A., 2006. Measuring the implications of sales and consumer inventory behavior. Econometrica. 74, 1637-1673. https://doi.org/10.1111/j.1468-0262.2006.00721.x.

Herrera, R., Clements, A.E., 2018. Point process models for extreme returns: harnessing implied volatility. Journal of Banking & Finance. 88, 161-175. https://doi.org/10.1016/j.jbankfin.2017.12.001.

Hull, C.L., 1932. The goal-gradient hypothesis and maze learning. Psychological Review. 39, 25-43. https://doi.org/10.1037/h0072640.

Hult, G.T.M., Sharma, P.N., Morgeson, F.V., Zhang, Y., 2019. Antecedents and consequences of customer satisfaction: Do they differ across online and offline purchases? Journal of Retailing. 95, 10-23. https://doi.org/10.1016/j.jretai.2018.10.003.

Hwang, Y., Gao, L., Mattila, A.S., 2020. What recovery options to offer for loyalty reward program members: Dollars vs. Miles? International Journal of Hospitality Management. 87, 102496. https://doi.org/10.1016/j.ijhm.2020.102496.

Jacobs, B., Fok, D., Donkers, B., 2021. Understanding large-scale dynamic purchase behavior. Marketing Science. 40, 844-870. https://doi.org/10.1287/mksc.2020.1279.

Jeong, J., Kim, D., Li, X., Li, Q., Choi, I., Kim, J., 2022. An empirical investigation of personalized recommendation and reward effect on customer behavior: A Stimulus-Organism-Response (SOR) model perspective. Sustainability. 14, 15369. https://doi.org/10.3390/su142215369.

Jones, M.A., Reynolds, K.E., Weun, S., Beatty, S.E., 2003. The product-specific nature of impulse buying tendency. Journal of Business Research. 56, 505-511. https://doi.org/10.1016/S0148-2963(01)00250-8.

Kaatz, C., 2020. Retail in my pocket– replicating and extending the construct of service quality into the mobile commerce context. Journal of Retailing and Consumer Services. 53, 101983. https://doi.org/10.1016/j.jretconser.2019.101983.

Karlis, D., Xekalaki, E., 2003. Choosing initial values for the EM algorithm for finite mixtures. Computational Statistics and Data Analysis. 41, 577-590. https://doi.org/10.1016/S0167-9473(02)00177-9.

Kim, B.D., Shi, M., Srinivasan, K., 2001. Reward programs and tacit collusion. Marketing Science. 20, 99-120. https://doi.org/10.1287/mksc.20.2.99.10191.

Kim, J.J., Steinhoff, L., Palmatier, R.W., 2021. An emerging theory of loyalty program dynamics. Journal of the Academy of Marketing Science. 49, 71-95. https://doi.org/10.1007/s11747-020-00719-1.

Kivetz, R., Simonson, I., 2003. The idiosyncratic fit heuristic: Effort advantage as a determinant of consumer response to loyalty programs. Journal of Marketing Research. 40, 454-467. https://doi.org/10.1509/jmkr.40.4.454.19383.

Kivetz, R., Urminsky, O., Zheng, Y., 2006. The goal-gradient hypothesis resurrected: Purchase acceleration, illusionary goal progress, and customer retention. Journal of Marketing Research. 43, 39-58. https://doi.org/10.1509/jmkr.43.1.39.

Kopalle, P.K., Sun, Y., Neslin, S.A., Sun, B., Swaminathan, V., 2012. The joint sales impact of frequency reward and customer tier components of loyalty programs. Marketing Science. 31, 216-235. https://doi.org/10.1287/mksc.1110.0687.

Koyama, S., Horie, T., Shinomoto, S., 2021. Estimating the time-varying reproduction number of COVID-19 with a state-space method. PLOS Computational Biology. 17. https://doi.org/10.1371/journal.pcbi.1008679.

Kuroda, M., 2020. EM algorithm. Kyoritsu Shuppan, Tokyo.

Kytö, E., Bult, H., Aarts, E., Wegman, J., Ruijschop, R.M., Mustonen, S., 2019. Comparison of explicit vs. implicit measurements in predicting food purchases. Food Quality and Preference. 78, 103733. https://doi.org/10.1016/j.foodqual.2019.103733.

Lee, K., Seo, B.K., 2017. Marked Hawkes process modeling of price dynamics and volatility estimation. Journal of Empirical Finance. 40, 174-200. https://doi.org/10.1016/j.jempfin.2016.08.004.

Lee, H.J., Choeh, J.Y., 2021. Motivations for obtaining and redeeming coupons from a coupon app: Customer value perspective. Journal of theoretical and applied electronic commerce research. 16, 22-33. https://doi.org/10.4067/S0718-18762021000200103.

Leenheera, J., Heerde, H.J., Bijmolt, T.H., Smidts, A., 2007. Do loyalty programs really enhance behavioral loyalty? An empirical analysis accounting for self-selecting members. International Journal of Research in Marketing. 24, 31-47. https://doi.org/10.1016/j.ijresmar.2006.10.005.

Lewis, E., Mohler, G., 2011. A nonparametric em algorithm for multiscale hawkes processes. Journal of Nonparametric Statistics.

Li, X., Dahana, W.D., Ye, Q., Peng, L., Zhou, J., 2021. How does shopping duration evolve and influence buying behavior? The role of marketing and shopping environment. Journal of Retailing and Consumer Services. 62, 102607. https://doi.org/10.1016/j.jretconser.2021.102607.

Li, L., Li, X., Qi, W., Zhang, Y., Yang, W., 2022. Targeted reminders of electronic coupons: Using predictive analytics to facilitate coupon marketing. Electronic Commerce Research. 22, 321-350. https://doi.org/10.1007/s10660-020-09405-4.

Liang, C.C., Lin, Y.W., 2023. Online promotion effects under time limitation - A study of survey and physiological signals. Decision Support Systems. 170, 113963. https://doi.org/10.1016/j.dss.2023.113963.

Liao, C., Lin, H.N., Luo, M.M., Chea, S., 2017. Factors influencing online shoppers' repurchase intentions: The roles of satisfaction and regret. Information & Management. 54, 651-668. http://dx.doi.org/10.1016/j.im.2016.12.005.

Lim, S.H., Lee, S., Kim, D.J., 2017. Is online consumers' impulsive buying beneficial for e-commerce companies? An empirical investigation of online consumers' past impulsive buying behaviors.

Information Systems Management. 34, 85-100. https://doi.org/10.1080/10580530.2017.1254458.

Liu, Y., 2007. The long-term impact of loyalty programs on consumer purchase behavior and loyalty. Journal of Marketing. 71, 19-35. https://doi.org/10.1509/jmkg.71.4.019.

Liu, Y., Sun, Y., Zhang, D., An analysis of "Buy X, Get One Free" reward programs. Operations Research. 69, 1823-1841. https://doi.org/10.1287/opre.2021.2128.

Lo, L.Y.S., Lin, S.W., Hsu, L.Y., 2016. Motivation for online impulse buying: A two-factor theory perspective. International Journal of Information Management. 36, 759-772. https://doi.org/10.1016/j.ijinfomgt.2016.04.012.

Ludwiczak, A., Osman, M., Jahanshahi, M., 2020. Redefining the relationship between effort and reward: Choice-execution model of effort-based decisions. Behavioural Brain Research. 383, 112474. https://doi.org/10.1016/j.bbr.2020.112474.

Madhavaram, S.R., Laverie, D.A., 2004. Exploring impulse purchasing on the Internet. Advances in Consumer Research. 31, 59-66.

Maggioni, I., 2016. What drives customer loyalty in grocery retail? Exploring shoppers' perceptions through associative networks. Journal of Retailing and Consumer Services. 33, 120-126. https://doi.org/10.1016/j.jretconser.2016.08.012.

Massar, S.A.A., Lim, J., Sasmita, K., Chee, M.W.L., 2016. Rewards boost sustained attention through higher effort: A value-based decision making approach. Biological Psychology. 120, 21-27. https://doi.org/10.1016/j.biopsycho.2016.07.019.

McCall, M., Voorhees, C., 2010. The drivers of loyalty program success: An organizing framework and research agenda. Cornell Hospitality Quarterly. 51, 35-52. https://doi.org/10.1177/1938965509355395.

McKenzie, J.E., Beller, E.M., Forbes, A.B., 2016. Introduction to systematic reviews and meta-analysis. Respirology. 21, 626-637. https://doi.org/10.1111/resp.12783.

Mercedes, M.P., Óscar, G.B., 2013. Studying motivations of store-loyal buyers across alternative measures of behavioural loyalty. European Management Journal. 31, 348-358. https://doi.org/10.1016/j.emj.2013.01.010.

Meyer-Waarden, L., 2015. Effects of loyalty program rewards on store loyalty. Journal of Retailing and Consumer Services. 24, 22-32. https://doi.org/10.1016/j.jretconser.2015.01.001.

Mirzaei, A., Baumann, C., Johnson, L.W., Gray, D., 2016. The impact of brand health on customer equity. Journal of Retailing and Consumer Services. 33, 8-16. https://doi.org/10.1016/j.jretconser.2016.07.001.

Mohler, G.O., Short, M.B., Brantingham, P.J., Schoenberg, F.P., Tita, G.E., 2011. Self-exciting point process modeling of crime. Journal of the American Statistical Association. 106, 100-108. https://doi.org/10.1198/jasa.2011.ap09546.

Mohler, G.O., 2013. Modeling and estimation of multi-source clustering in crime and security data. The Annals of Applied Statistics. 7, 1525-1539. https://doi.org/10.1214/13-AOAS647.

Monsuwé, T.P.Y., Dellaert, B.G.C., Ruyter, K., 2004. What drives consumers to shop online? A literature review. International Journal of Service Industry Management. 15, 102-121. https://doi.org/10.1108/09564230410523358.

Morisada, M., Miwa, Y., Dahana, W.D., 2019. Identifying valuable customer segments in online fashion markets: An implication for customer tier programs. Electronic Commerce Research and Applications. 33, 100822. https://doi.org/10.1016/j.elerap.2018.100822.

Mowrey, C.H., Parikh, P.J., Gue, K.R., 2018. A model to optimize rack layout in a retail store. European Journal of Operational Research. 271, 1100-1112. https://doi.org/10.1016/j.ejor.2018.05.062.

Mutius, B.V., Huchzermeier, A., 2021. Customized targeting strategies for category coupons to maximize CLV and minimize cost. Journal of Retailing. 97, 764-779. https://doi.org/10.1016/j.jretai.2021.01.004.

Nastasoiu, A., Bendle, N.T., Bagga, C.K., Vandenbosch, M., Navarro, S., 2021. Separating customer heterogeneity, points pressure and rewarded behavior to assess a retail loyalty program. Journal of the Academy of Marketing Science. 49, 1132-1150. https://doi.org/10.1007/s11747-021-00782-2.

Nishikawa, H., Schreier, M., Fuchs, C., Ogawa, S., 2017. The value of marketing crowdsourced new products as such: Evidence from two randomized field experiments. Journal of Marketing Research. 54, 525-539. https://doi.org/10.1509/jmr.15.0244.

Nunes, J.C., Drèze, X., 2011. Recurring goals and learning: the impact of successful reward attainment on purchase behavior. Journal of Marketing Research. 48, 268-281. https://doi.org/10.1509/jmkr.48.2.268.

Ogata, Y., 1981. On Lewis' simulation method for point processes. IEEE Transactions on Information Theory. 27, 23-31. https://doi.org/10.1109/TIT.1981.1056305.

Ogata, Y., 1988. Statistical models for earthquake occurrences and residual analysis for point processes. Journal of the American Statistical Association. 83, 9-27. https://doi.org/10.2307/2288914.

Ogata, Y., 1998. Space-time point-process models for earthquake occurrences. Annals of the Institute of Statistical Mathematics. 50, 379-402. https://doi.org/10.1023/A:1003403601725.

Ogata, Y., 2017. Statistics of earthquake activity: models and methods for earthquake predictability studies. Annual Review of Earth and Planetary Sciences. 45, 497-527. https://doi.org/10.1146/annurev-earth-063016-015918.

Omi, T., Nomura, S., 2019. Time Series Analysis for Point Processes. Kyoritsu Pub, Tokyo.

Palmeira, M., Pontes, N., Thomas, D., Krishnan, S., 2016. Framing as status or benefits? Consumers' reactions to hierarchical loyalty program communication. European Journal of Marketing. 50, 488-508. https://doi.org/10.1108/EJM-02-2014-0116.

Park, C.H., Park, Y.H., Schweidel, D.A., 2018. The effects of mobile promotions on customer purchase dynamics. International Journal of Research in Marketing. 35, 453-470. https://doi.org/10.1016/j.ijresmar.2018.05.001.

Park, C.H., Yoon, T.J., 2022. The dark side of up-selling promotions: Evidence from an analysis of cross-brand purchase behavior. Journal of Retailing. 98, 647-666. https://doi.org/10.1016/j.jretai.2022.03.005.

Praag, B.M.S.V., Vermeulen, E.M., 1993. A count-amount model with endogenous recording of observations. Journal of Applied Econometrics. 8, 383-395. https://doi.org/10.1002/jae.3950080406.

Radder, L., Eyk, M.V., Swiegelaar, C., 2015. Levels of customer loyalty and perceptions of loyalty programme benefits: a south african retail example. The Retail and Marketing Review. 11, 92-105. https://journals.co.za/doi/10.10520/EJC177067.

Reinartz, W.J., Kumar, V., 2000. On the profitability of long-life customers in a noncontractual setting: An empirical investigation and implications for marketing. Journal of Marketing. 64, 17-35. https://doi.org/10.1509/jmkg.64.4.17.18077.

Reinhart, A., 2018. A review of self-exciting spatio-temporal point processes and their applications. Statistical Science. 33, 299-318. https://doi.org/10.1214/17-STS629.

Rosenbaum, P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. Biometrika. 70, 41-55. https://doi.org/10.1093/biomet/70.1.41.

Rossi, P.E., McCulloch, R.E., Allenby, G.M., 1996. The value of purchase history data in target marketing. Marketing Science. 15, 321-340. https://doi.org/10.1287/mksc.15.4.321.

Rubin, D.B., 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. Journal of Educational Psychology. 66, 688-701. https://doi.org/10.1037/h0037350.

Rubin, D.B., 1978. Bayesian inference for causal effects: The role of randomization. The Annals of Statistics. 6, 34-58. https://doi.org/10.1214/aos/1176344064.

Schoenberg, F.P., 2013. Facilitated estimation of ETAS. Bulletin of the Seismological Society of America. 103, 601-605. https://doi.org/10.1785/0120120146.

Septianto, F., An, J., Chiew, T.M., Paramita, W., Tanudharma, I., 2019. The similar versus divergent effects of pride and happiness on the effectiveness of loyalty programs. Journal of Business Research. 99, 12-22. https://doi.org/10.1016/j.jbusres.2019.02.021.

Sharma, D., Verma, V., 2014. Psychological and economic considerations of rewards programs. Journal of Retailing and Consumer Services. 21, 924-932. https://doi.org/10.1016/j.jretconser.2014.08.010.

Sharp, B., Sharp, A., 1997. Loyalty programs and their impact on repeat-purchase loyalty patterns. International Journal of Research in Marketing. 14, 473-486. https://doi.org/10.1016/S0167-8116(97)00022-0.

Song, P., Wang, Q., Liu, H., Li, Q., 2020. The value of buy-online-and-pickup-in-store in omnichannel: Evidence from customer usage data. Production and Operations Management. 29, 995-1010. https://doi.org/10.1111/poms.13146.

Steiner, P.M., Cook, T.D., Shadish, W.R., Clark, M.H., 2010. The importance of covariate selection in controlling for selection bias in observational studies. Psychological Methods. 15, 250-267. https://doi.org/10.1037/a0018719.

Su, Y., Lee, C., Lovelock, B., 2023. Airline and hotel loyalty programme diversity and choice: Effects of personality, cultural, and socio-demographic factors. Journal of Travel & Tourism Marketing. 40, 653-677. https://doi.org/10.1080/10548408.2023.2285926.

Sundén, A.E., Surette, B.J., 1998. Gender differences in the allocation of assets in retirement savings plans. The American Economic Review. 88, 207-211. https://www.jstor.org/stable/116920.

Tanford, S., Raab, C., Kim, Y.S., 2011. The influence of reward program membership and commitment on hotel loyalty. Journal of Hospitality & Tourism Research. 35, 279-307. https://doi.org/10.1177/1096348010382236.

Tanford, S., 2013. The impact of tier level on attitudinal and behavioral loyalty of hotel reward program members. International Journal of Hospitality Management. 34, 285-294. https://doi.org/10.1016/j.ijhm.2013.04.006.

Tanford, S., Baloglu, S., 2013. Applying the loyalty matrix to evaluate casino loyalty programs. Cornell Hospitality Quarterly. 54, 333-346. https://doi.org/10.1177/1938965512464694.

Tanford, S., Malek, K., 2015. Segmentation of reward program members to increase customer loyalty: the role of attitudes towards green hotel practices. Journal of Hospitality Marketing & Management. 24, 314-343. https://doi.org/10.1080/19368623.2014.907759.

Taylor, G.A., Neslin, S.A., 2005. The current and future sales impact of a retail frequency reward program. Journal of Retailing. 81, 293-305. https://doi.org/10.1016/j.jretai.2004.11.004.

Ton, Z., Raman, A., 2010. The effect of product variety and inventory levels on retail store sales: a longitudinal study. Production and Operations Management. 19, 546-560. https://doi.org/10.1111/j.1937-5956.2010.01120.x.

Veen, A., Schoenberg, F.P., 2008. Estimation of space-time branching process models in seismology using an EM-type algorithm. Journal of the American Statistical Association. 103, 614-624. https://doi.org/10.1198/016214508000000148.

Wang, R.J.H., Malthouse, E.C., Krishnamurthi, L., 2015. On the go: How mobile shopping affects customer purchase behavior. Journal of Retailing. 91, 217-234. https://doi.org/10.1016/j.jretai.2015.01.002.

Weisberg, J., Te'eni, D., Arman, L., 2011. Past purchase and intention to purchase in e-commerce: The mediation of social presence and trust. Internet Research. 21, 82-96. https://doi.org/10.1108/10662241111104893.

Wierich, R., Zielke, S., 2014. How retailer coupons increase attitudinal loyalty - The impact of three coupon design elements. European Journal of Marketing. 48, 699-721. https://doi.org/10.1108/EJM-01-2012-0030.

Wu, J., Huang, L., Zhao, J.L., Hua, Z., 2015. The deeper, the better? Effect of online brand community activity on customer purchase frequency. Information and Management. 52, 813-823. https://doi.org/10.1016/j.im.2015.06.001.

Xu, Y., Huang, J.S., 2014. Effects of price discounts and bonus packs on online impulse buying. Social Behavior and Personality. 42, 1293-1302. https://doi.org/10.2224/sbp.2014.42.8.1293.

Xu, L., Duan, J.A., Whinston, A., 2014. Path to purchase: a mutually exciting point process model for online advertising and conversion. Management Science. 60, 1392-1412. https://doi.org/10.1287/mnsc.2014.1952.

Yamada, H., Sato, T., 2021. An analysis of mechanism for customers' purchase amount and number of visits in department store. Journal of the Operations Research Society of Japan. 64, 12-30. https://doi.org/10.15807/jorsj.64.12.

Yang, S.Y., Liu, A., Chen, J., Hawkes, A., 2017. Applications of a multivariate Hawkes process to joint modeling of sentiment and market return events. Quantitative Finance. 18, 295-310. https://doi.org/10.1080/14697688.2017.1403156.

Zeithaml, V.A., Rust, R.T., Lemon, K.N., 2001. The customer pyramid: creating and serving profitable customers. California Management Review. 43, 118-142. https://doi.org/10.2307/41166104.

Zhang, H., Yu, Y., Qin, Y., 2023. The effects of constrained mobile coupons in the mobile channel. Journal of Retailing and Consumer Services. 75, 103542. https://doi.org/10.1016/j.jretconser.2023.103542.

Zhao, Y., Li, Y., Wang, N., Zhou, R., Luo, X., 2022. A meta-analysis of online impulsive buying and the moderating effect of economic development level. Information Systems Frontiers. 24, 1667-1688. https://doi.org/10.1007/s10796-021-10170-4.

Zhuang, J., Ogata, Y., Jones, D., 2002. Stochastic declustering of space-time earthquake occurrences. Journal of the American Statistical Association. 97, 369-380. https://doi.org/10.1198/016214502760046925.