Timing customer reactivation initiatives

Holtrop, Niels; Wieringa, Jaap E.

Published in:
International Journal of Research in Marketing

DOI:
10.1016/j.ijresmar.2023.05.001

Publication date:
2023

International Journal of Research in Marketing 40 (2023) 570–589

Contents lists available at ScienceDirect

International Journal of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Timing customer reactivation initiatives

Niels Holtrop[a,*], Jaap E. Wieringa[b]

[a] School of Business and Economics, Maastricht University, Tongersestraat 53, 6211 LM Maastricht, The Netherlands
[b] Department of Marketing, Faculty of Economics and Business, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands

ARTICLE INFO

Article history:
First received on 22 June 2020 and was under review for 13 months
Available online 29 May 2023

Area Editor: Lan Luo
Accepting Editor: Renana Peres

Keywords:
Customer reactivation
Control chart
Customer base analysis
Field test
Non-contractual setting

ABSTRACT

Firms operating in non-contractual settings apply customer reactivation initiatives such as email messages to stimulate customers who have become inactive temporarily or permanently to resume their transaction activities. Thus, firms need to know *which* customers are inactive, and *when* a customer becomes inactive. Existing approaches struggle to distinguish active from inactive customers and do not provide time-scale estimates of when to send reactivation mails. To address these shortcomings, we develop an approach to target and time the sending of reactivation mails. Building on control chart methods, we introduce a gamma–gamma control chart, modelling the average customer interpurchase time and the variation therein to determine activity boundaries. Crossing these boundaries signals a potential change in a customer's purchasing activity, providing a signal to initiate customer reactivation. A field experiment in the greetings and gifts industry, supported by several additional analyses, illustrates the improved performance of our approach when it comes to signaling customer activity against a wide range of competing models. The improved performance of our method occurs particularly in settings where customers vary strongly in purchase and inactivity patterns.

1. Introduction

Many organizations operating in non-contractual settings (e.g., retailing, catalogs, charity donations) proactively communicate with customers who reduced their transaction frequency temporarily or stopped their transactions entirely. Such proactive communication, termed *customer reactivation* (e.g., Blömeke, Clement, & Bijmolt, 2010), is addressed frequently in professional publications (Pokornyik, 2017; Stevens, 2017). However, academic guidance on customer reactivation is limited.

Customer reactivation requires identifying *which* customers are active and inactive[1]. However, a customer's true activity state cannot be directly observed due to the non-contractual setting. Beyond potential inactivity, customer reactivation also requires knowledge of *when* the customers changed their transaction levels. Given that customers can transact at any time with the organization, are heterogeneous in their transaction frequencies, and do not transact at set time intervals, detecting such a change in transaction patterns is not straightforward. In this paper, we address these questions with the aim of implementing effective customer reactivation initiatives.

* Corresponding author.
E-mail addresses: n.holtrop@maastrichtuniversity.nl (N. Holtrop), J.E.Wieringa@rug.nl (J.E. Wieringa).

[1] Throughout the paper, we use the term "inactive customers," although the non-contractual setting does not allow us to distinguish whether these customers have ceased activity temporarily (i.e., are inactive) or permanently (i.e., have churned). Our aim is to identify customers who deviate from their "normal" purchasing state, which does not limit our goals and findings.

https://doi.org/10.1016/j.ijresmar.2023.05.001
0167-8116/

The implementation of customer reactivation initiatives depends on 1) identifying inactive customers, 2) determining the time they reduced their activities, and 3) initiating communication with the customer. Since e-mail is used in 90% of customer reactivation initiatives (next to e.g., phone calls and physical mailings)[2], we focus on communication through mail in this paper. The identification of inactive customers in non-contractual settings has been studied extensively. Accordingly, model-based approaches are common, giving rise to stochastic latent attrition models, such as the Pareto/NBD model (Schmittlein, Morrison, & Columbo, 1987) and its variants (e.g., the BG/NBD model, see Fader, Hardie, & Lee, 2005a). However, Wübben and Von Wangenheim (2008) find that such models perform poorly when predicting individual-level future transactions, which is required for determining customer-level inactivity. The authors instead recommend the use of business rules, such as the time-since-last-transaction (hiatus) heuristic, where a customer is assumed to have churned after × days/weeks/-months without transacting.

Research is scarcer regarding the time at which customers become inactive, as well as the consequences of sending a mailing in response. Studies have acknowledged the importance of intervention timing, as prolonged reduced activity adversely affects recency and decreases future transaction likelihood (Bult & Wansbeek, 1995; Fader, Hardie, & Lee, 2005b; Neslin et al., 2013). Furthermore, the time between successive mailings affects the overall customer equity (Drèze & Bonfrer, 2008). Finally, the literature suggests that guidance on intervention timing should provide an exact time (i.e., day/week/month) or calendar date rather than the activity probabilities produced by stochastic models, as managers make decisions—including those on customer reactivation—using exact time scales (e.g., Korkmaz, Kuik, & Fok, 2013). Indeed, the managerial question is *when* an action should be taken. For marketing scheduling purposes, this is more easily communicated in terms of a calendar date (e.g., January 5) rather than a probability (e.g., 0.63). Given that mailing actions are implemented at short notice, such estimates should be readily available as well. Latent attrition models do not provide exact time-scale estimates, as they focus on detecting if defections occur rather than when they occur. The question then arises on how to provide such time-scale estimates linked to the moment at which a customer becomes inactive and on whether sending a mailing in response to this event can restore customers' transaction frequencies. We develop an approach that generates such time-scale estimates.

Our approach relies on techniques developed in the statistical quality control literature (Montgomery, 2009). This stream of research is concerned with monitoring and controlling (industrial) processes in the face of process variability, making it ideally suited to our goals. In particular, we focus on the so-called control charts (Shewhart, 1931), which are used to monitor a process variable (e.g., the average concentration of tin in a chemical bath) and detect situations where the process does not adhere to its requirements (Wieringa, 1997). We fuse this control-chart approach with existing models for purchasing in non-contractual settings (e.g., Colombo & Jiang, 1999; Fader, Hardie, & Lee, 2005b) to develop a customer-monitoring system that provides frequent updates on the (in)activity of customers at the time scale. The system provides a signal when a customer's transaction activity is deemed to be reduced given his/her regular transaction pattern, which is used as an indicator to send a reactivation mail. Thus, we take advantage of the latent attrition models' ability to detect defections while addressing their limitation of not providing information on when defection occurs.

We demonstrate the validity of our approach using empirical and simulation studies. A field test in the greetings and gifts industry demonstrates the efficacy of our approach in a real-life setting. We find a 1.9-percentage-point increase in customer activity compared with the current firm policy, as well as a 3.5-percentage-point increase compared with the control group (no mailing received). Using the field test data, we subsequently use offline policy analysis (e.g., Hitsch & Misra, 2018) to compare our approach with a range of competing models identified from the literature. We find that the control-chart approach compares favorably to the competing methods in terms of determining customer activity, including recent causal machine-learning methods (Cui et al., 2020; Wager & Athey, 2018). Finally, we investigate the performance of our approach for different customer segments using a simulation approach[3]. We find that the improved performance of our approach relative to a set of benchmarks arises in situations characterized by heterogeneous purchase and/or inactivity processes, whereas the benchmark models fare better in settings where purchase and inactivity times do not differ strongly between customers. Furthermore, improved performance arises for longer ($\geq$52 weeks) calibration and prediction periods; performance is equivalent to the benchmarks for shorter periods. Based on these findings, we provide managerial guidance to effectively implement customer reactivation.

The remainder of this paper is organized as follows. In Section 2, we review the existing literature related to customer reactivation. Section 3 introduces the control-chart method and outlines the development of our approach. We also introduce a mixture version of our approach to account for time-varying transaction states. Section 4 introduces the data and the empirical setting for our study, while Section 5 provides the results of our analyses. We conclude in Section 6 with a discussion of the implications of our work and provide directions for future research.

2. Prior studies related to the customer reactivation process

Customer reactivation aims to motivate customers who have reduced or ceased transacting to resume their transactions by sending them mailing (Blömeke et al., 2010). The successful implementation of customer reactivation initiatives in a non-

[2] https://www.targetmarketingmag.com/article/2016-customer-acquisition-retention-and-the-best-roi/.

[3] We thank the two anonymous reviewers for suggesting the offline policy analysis and the simulation approach.

contractual setting requires firms to 1) identify which customers have become inactive, 2) determine the specific time the inactivity occurred, and 3) contact these customers with a reactivation mailing at that point in time. Our work is informed by two marketing research streams: 1) research focusing on identifying *which* customers should be approached with a marketing incentive and 2) studies focusing on *when* to target customers (Ascarza et al., 2018). The main insights from these literature streams that inform our approach are discussed below. We combine insights from these research streams with techniques from the statistical quality control literature (i.e., control charts) in Section 3 and show how the control charts can help develop a method for customer reactivation.

2.1. Studies determining whom to target in non-contractual settings

The challenge of which customers to target at what time has been investigated across different studies (see Table 1). Two main underlying literature streams emerge here, namely, 1) studies in the direct mailing domain and 2) studies in the customer management domain. Determining which customers to target is an important challenge in the direct mailing domain (Bult & Wansbeek, 1995). For example, studies have shown that consideration of customers' mailing and transaction histories is important when determining which customers to approach with mailings (e.g., Gönül & Shi, 1998; Gönül, Kim & Shi, 2000; Gönül & Ter Hofstede, 2006; Simester, Sun, & Tsitsiklis, 2006), as they may influence the effectiveness of such actions (Van Diepen, Donkers, & Franses, 2009). Our approach takes this information into account as well. The main aim of these existing studies is to enhance overall profitability by maximizing the benefits of increased transaction value minus the cost of sending mailings.

Most studies offer recommendations about mailing volume, such as the optimal number of mailings (e.g., Neslin et al., 2013; Van Diepen, Donkers, & Franses, 2009), and ignore when to send these mailings. However, two exceptions exist. Drèze and Bonfrer (2008) show at an aggregate level that the time between mailings affects the overall customer equity, while (Gönül et al., 2000) use segment-specific timing estimates to determine when to send a mailing. Nevertheless, both approaches do not inform managers faced with scheduling these mailings about when to send them (i.e., the calendar date) to *individual* customers, which is our objective in this paper.

The customer management domain has more extensively studied the question of who (and when) to target customers. Existing models, mainly of the latent attrition/buy-till-you-die type (BTYD, Fader, Hardie, & Lee, 2005a), provide the probability of making future transactions or *P(Alive)*. *P(Alive)* is useful for customer base valuation (e.g., Fader, Hardie, & Lee, 2005) and can guide which customers to target, but these latent attrition models do not provide information on when to target such customers. The fact that it is a probability instead of a deterministic number makes it difficult to translate into a binary active/inactive decision required for determining when a customer should be targeted. Prior work thus suggests 1) determining a cut-off value for *P(Alive)* to aid the translation to a binary decision (Reinartz & Kumar, 2000; Wübben & Von Wangenheim, 2008) or 2) estimating a reactivation probability with an appropriate cut-off (Ma, Tan, & Shu, 2015).

However, Wübben and Von Wangenheim (2008) show that such cut-off approaches only capture *aggregate* behavior accurately, making them less suited for our intended goal of identifying *individual* customers who changed their purchasing activity level. Korkmaz et al. (2013) recognize this problem and extend several BTYD models to predict the next individual purchase time.

2.2. Studies determining when to target in non-contractual settings

For reactivation purposes, it is not only important to determine *which* customers are less committed but also *when* they change their transaction behavior (Neslin et al., 2013). While cut-off-based methods provide this information when a customer's probability is below the cut-off value, continuous model updates are required to produce up-to-date forecasts to detect the moment at which the cut-off value is reached. The approach of Korkmaz et al. (2013) suffers from a similar limitation. Furthermore, given that past literature has mainly focused on mailings with substantial production costs (physical mail, catalogs, and phone calls), our focus on e-mail shifts the focus from *whom* to target to *when* to target as an important criterion, as the production costs of approaching the wrong customer are negligible in monetary terms. Revenues generated by reactivation campaigns are then determined by customer response, and firms would want to avoid over- or undersending reactivation mailings in this scenario to avoid irritation or lost opportunities.

Alternative techniques for accurately determining customer inactivity have been explored, aimed at modeling the time between transactions (interpurchase time or IPT). Allenby, Leone, and Jen (1999) pursue these alternatives in the direct mailing domain, while Platzer and Reutterer (2016) and Reutterer, Platzer, and Schröder (2021) do so in the customer management domain. The former study uses a generalized gamma model to model purchase timing but does not use this information to guide targeting decisions. The two latter studies examine the concept of purchase timing regularity, showing that including purchase timing using IPT provides additional information to improve the performance of the stochastic purchase models. To this end, these papers introduce the Pareto-GGG and the MBG/CNBD-k models, respectively, but do not consider the question of when customers should be targeted. Building on these former insights, we develop a reactivation model that uses IPT as the underlying measure. The advantage of considering IPT (measured as days/weeks/months) is that cut-off transformations need not be applied to our model outcomes, as the model produces estimates at the time scale rather than probabilistic predictions (e.g., Allenby, Leone, & Jen, 1999). Expanding these prior studies, our approach provides insights on whom to target and at what time.

Table 1
Selected prior studies related to the customer reactivation process.

	Focus on when to target	Focus on whom to target	Customer activity as outcome	Profit as outcome	Field test used	Domain
Allenby, Leone, & Jen (1999)	✓					Direct mail
Ascarza & Hardie (2013)		✓	✓			Customer management
Ascarza (2018)		✓	✓		✓	Customer management
Blömeke et al. (2010)		✓	✓		✓	Direct mail
Bult & Wansbeek (1995)		✓		✓		Direct mail
Drèze & Bonfrer (2008)	✓		✓			Direct mail
Fader, Hardie, & Lee (2005a)		✓				Customer management
Gönül & Ter Hofstede (2006)		✓		✓		Direct mail
Gönül et al. (2000)	✓	✓	✓	✓		Direct mail
Gönül & Shi (1998)		✓		✓		Direct mail
Korkmaz et al. (2013)	✓	✓	✓			Customer management
Lemmens & Gupta (2020)		✓		✓	✓	Customer management
Ma et al. (2015)		✓	✓			Customer management
Neslin et al. (2013)		✓		✓		Direct mail
Platzer & Reutterer (2016)		✓				Customer management
Reinartz & Kumar (2000)		✓		✓		Customer management
Reutterer, Platzer, & Schröder (2021)		✓				Customer management
Schmittlein et al. (1987)		✓				Customer management
Seetharaman & Chintagunta (2003)			✓			Customer management
Simester et al. (2006)		✓		✓	✓	Direct mail
Van Diepen et al. (2009)		✓		✓	✓	Direct mail
Wübben & Von Wangenheim (2008)		✓	✓			Customer management
This study	✓	✓	✓	✓	✓	Customer management

Time-scale predictions can also be generated without relying directly on IPT. Seetharaman and Chintagunta (2003) show that the proportional hazards model (PHM), which accounts for heterogeneity, provides a good way to model transaction event timing. However, they do not compare their models to those previously discussed but only to other PHMs, making their relative performance unknown. Second, following Korkmaz et al. (2013), the Pareto/GGG and MBG/CNBD-k models can be adapted for reactivation timing by predicting the next transaction time, but they have not been used for this purpose thus far. Third, hidden Markov models (HMM, e.g., Ascarza & Hardie, 2013) can also provide time-scaled estimates. However, to our knowledge, this has not been pursued previously in the customer management setting. We include all these models in our comparison in Section 5.3 to identify their ability to provide accurate transaction timing in our setting.

We note that all the methods discussed thus far (including our own approach) form the basis of the intervention (here: mailing) decision on the *risk* of a customer: What is the likelihood that a customer turns inactive/defects regardless of any targeted firm intervention? However, recent studies (Ascarza, 2018; Lemmens & Gupta, 2020) argue that intervention decisions should be based on *lift*: What is the difference in response due to an intervention? Lift estimates can be obtained by combining field experimental data with machine learning approaches. Ascarza (2018) shows that targeting lift is more effective in reducing customer churn, while Lemmens and Gupta (2020) show that combining lift with a profit-based loss function can yield superior profits. While these prior studies investigate the impact of which customers to target (i.e., those with the highest lift), they do not investigate whether the timing of the intervention matters. Nevertheless, we compare our approach to the lift-based approaches in Section 5.3.

3. Developing a control chart method for customer reactivation

The prior section highlights the existing knowledge and challenges related to determining customer activity. In this section, we introduce the control charts from the statistical quality control literature as a means to monitor customer behavior over time and argue that control charts can be useful tools for determining customer (in)activity. Subsequently, we discuss the development of our control chart approach for customer reactivation.

3.1. Statistical quality control and control charts

Determining a suitable time at which individual customers should be approached with a reactivation mailing requires an inactivity signal for *each* customer at the/any point in time when the reactivation decision is made, which can be daily, weekly (as in our field test), biweekly, or monthly. Given the typically large number of customers in a customer base, this process requires continuously *monitoring* a substantial number of units (i.e., customers). Monitoring numerous units is the focus of the statistical quality control literature.

This stream of literature is concerned with monitoring industrial processes and intervening when disturbances are detected to ensure their continuation (Montgomery, 2009). Statistical quality control approaches are used to monitor inventory stocks (Ernst, Guerro, & Roshwalb, 1993), wafer stepper production (Does et al., 1999), and the tin-plating of surface-mounted diodes (Wieringa, 1997). Marketing applications include the selection of marketing test panel members (Marcuse, 1945), monitoring market shares and promotions (Crespy Stearns, & Krebhiel, 1995), and yearly monitoring of aggregate customer satisfaction scores (Sharma, Niedrich, & Dobbins, 1999). Different from these prior marketing studies, we monitor the transaction behavior of several individual customers near-continuously instead of monitoring the aggregate-level variables at limited time points (e.g., monthly, yearly), resulting in individual-level predictions of future transaction behavior.

In the statistical quality control literature, various methods have been developed to efficiently monitor several units simultaneously. One of the earliest and most prominent examples is the control chart (Shewhart, 1931). A control chart monitors the performance of a process through a target variable, such as the average concentration of a chemical substance. By measuring this target variable at different points in time and plotting the resulting time series, a chart is created. The chart also includes predetermined bounds (control limits) within which the process is allowed to fluctuate. If the target variable crosses one of these bounds, a signal is generated, and an intervention by the process owner can bring the process back within its bounds. A fictitious example control chart is shown in Fig. 1[4]. Situations wherein one of the bounds is crossed indicate that the process is not performing according to normal operations; thus, it is said to be *out of control*. This excess variation is due to what Shewhart (1931) calls special causes (e.g., the chemical concentration deviates strongly from the mean (target) concentration due to some external contamination).

However, some variability is allowed, given that each process suffers from normal variation that is inherent to the process (common causes; Shewhart, 1931). The bounds of the control chart are determined such that special causes can be distinguished from common causes of variation. Action is only required if the presence of a special cause of variation is detected by the control chart. More importantly, one should not interfere with the process as long as it is in control (i.e., moves within the boundaries determined by the variation due to common causes). Doing so would only lead to additional variation and potentially destabilize the process ("tampering with the process"; Deming, 1982).

3.2. Using control charts for customer reactivation

Control charts can also be used to monitor the purchasing behavior of individual customers, providing a suitable solution for our problem of deciding which customers to approach with a reactivation action. By design, control charts intend to monitor a variable over time. Adapting this to the problem of determining the changes in a customer's activity level, we track the IPTs. A time series of IPT observations are formed by tracking this variable at the individual customer level across transaction occasions.

Thus, we derive customer-specific trajectories of transaction behavior with associated bandwidths around them. Using these bandwidths, we can convert the IPT from a time-scale measure to a binary active/inactive decision. This conversion occurs at the point where the time since the customer's last transaction exceeds the bandwidth of that customer's normal transaction behavior. That is, if one of the control chart limits is crossed, indicating an increased probability of a special cause for that customer, we consider this a signal that firms can use to initiate an intervention. Earlier intervention is not advisable in this case because a transaction could have been delayed due to common cause variation in transaction behavior.

This separation of signals is important for effective customer management. As Ascarza, Iyengar, and Schleicher (2016) show, firms should not react to all changes in customer behavior. They find that recommending mobile phone price plans to decrease customer churn can actually increase churn. In their application, the adverse effect of the recommendations is caused by making customers who are unaware (not inactive in our setting) become aware of their unsuitable phone plans. This is an example of "tampering with the process" (Deming, 1982), which may have negative consequences (e.g., customer churn). The control chart approach addresses this situation by separating common and special causes of variation. Firms should not interfere in a stable process and only target customers who change their normal behaviors. Interventions initiated

[4] An important difference between this fictitious example and the control chart we develop is that our control chart boundaries are allowed to change over time, which is not the case for the standard control chart used in the statistics quality control theory. Sections 3.3 and 3.4 discuss how this is achieved, while Section 5.1 and Fig. 3 illustrate our control charts.

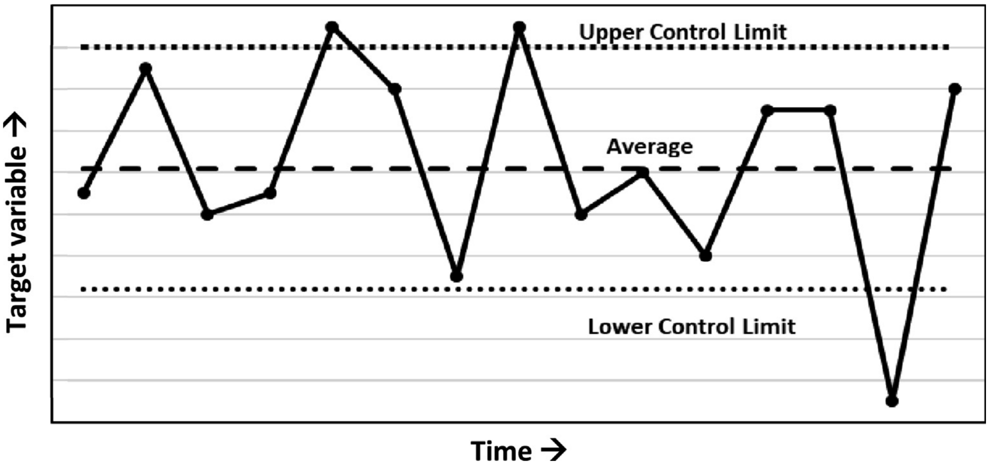

Fig. 1. A fictitious control chart with fixed upper and lower control limits.

before a plausible inactivity signal is present will not only target inactive customers but also make them either ineffective or even cause adverse effects.

3.3. Developing a control-chart method for customer reactivation timing

Given our objective of developing a control-chart approach for reactivation purposes, we need to select a variable that characterizes a customer's transaction behavior. The control chart then monitors this variable to guide the timing of reactivation actions (Montgomery, 2009). Based on our discussion in Section 2.1, we use IPT as a measure of transaction timing.

We develop our model by first considering the characteristics of the statistical process underlying our variable of interest for the situation when the customer is active. Our model development is based on the following assumptions, adapted from Fader, Hardie, and Lee (2005b):

- For a given transaction occasion, the time since the last transaction varies around the average customer-specific IPT.
- The average IPT varies across customers but not over time for a specific customer (given that we model the customer's normal purchasing behavior affected by common causes only).

The latter assumption implies that there is a true process mean to be estimated, with noisy variation around the mean. We relax this assumption in Section 3.6, where we consider the impact of "hot" and "cold" transaction states. To accurately model the time between transactions, the properties of the process distribution (i.e., IPT in our case) must be considered. In Fig. W3a in the Web Appendix, we summarize this distribution based on the data described in detail in Section 4.1.2. The distribution appears to be right-skewed, and the time between transactions is always non-negative. These characteristics indicate that a normal distribution is not suitable to model IPT. Thus, we adopt the gamma distribution to model this process (e.g., Platzer & Reutterer, 2016; Zhang et al., 2007). Another attractive feature of the gamma distribution is that it relaxes the assumption of a Poisson-distributed transaction process (with exponential IPTs) to be non-Poisson (Platzer & Reutterer, 2016; Reutterer, Platzer & Schröder, 2021), which accounts for irregularity in transaction timing.

This assumption would have been sufficient when considering the transaction process of a single customer. However, given that we are seeking to model the transaction processes of several customers, it seems unreasonable to assume a similar distribution for each customer (see Assumption 2 above). We assume that heterogeneity in transaction processes exists across the population (e.g. customers have different expected IPTs) and that this can be modeled with another gamma distribution. We thus follow Fader, Hardie, and Lee (2005b) and use the gamma–gamma model of Colombo and Jiang (1999). One advantage of this is that we can obtain closed-form solutions for our key expressions later[5].

To develop a control chart that can monitor the IPT process, we first consider the general form of such a chart. In this study, we focus on the simplest form following Shewhart (1931). This control chart plots the individual observations along with an upper control limit (Montgomery, 2009), which we can write as [6]

$$CC(x_i) = \mu(x_i) + c \times \sigma(x_i) \tag{1}$$

[5] We validate the gamma–gamma assumption for our empirical data and find that this distribution shows a good fit to the distribution of the data presented in Table 3. The results are provided in the Web Appendix W3.

[6] Control charts also have a lower bound (see Fig. 1), defined as $CC(x_i) = \mu(x_i) - c \times \sigma(x_i)$. However, as purchase time is naturally bounded at 0, we only consider charts with an upper control limit.

In this equation, μ and σ can be replaced with suitable estimators of this quantity depending on the underlying process distribution, while x_i represents the variable monitored in the control chart of customer i (IPT in our case). The width of the control chart c is often taken as 3 based on the assumption that successive observations of the process are identically and normally distributed. Under this assumption, this condition corresponds to a false alarm probability (i.e., giving a signal when no signal should be given) of 1 – 0.9973 = 0.0027, that is, false alarms are highly unlikely, and a signal from a control chart should be a cause for concern (Shewhart, 1931). Given our earlier exposition on the non-normality of our underlying distribution, we will replace each of these quantities with forms suitable to our desired application. Thus, we derive expressions for μ and σ based on the gamma–gamma model to deal with non-normality and propose a procedure to determine c, given this non-normal data structure.

3.4. Deriving $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ for the gamma–gamma model

We derive expressions for the mean and the standard deviation of a process that can be characterized by a gamma–gamma model. We use these expressions to determine the average IPT of a customer after x_i transactions, as well as the variation therein, which is needed for calculating the control limits of the chart.

Following Fader and Hardie (2013), let n_i denote the number of transactions of a customer i and let $T_{i1}, T_{i2}, \cdots, T_{in_i}$ denote the time between each transaction[7]. Define $\bar{T}_i = \sum_{j=1}^{n_i} \frac{T_{ij}}{n_i}$ as an estimate of the (unobserved) true average IPT of a customer i, denoted as ξ_i. We are interested in two quantities related to ξ_i: its conditional mean denoted as $\mathrm{E}\left(T_{ij}|\bar{T}_i = \bar{t}_i, n_i\right)$ and its conditional variance denoted as $\mathrm{Var}\left(T_{ij}|\bar{T}_i = \bar{t}_i, n_i\right)$. These two quantities will serve as estimates for μ and σ. To arrive at these expressions, we formalize the model of Colombo and Jiang (1999) as follows:

1. We assume that $T_{ij} \sim gamma(p, v_i)$ with shape parameter p and rate parameter v_i. This implies $E(T_{ij}|p, v_i) = \xi_i = \frac{p}{v_i}$, which gives $\bar{T}_i \sim gamma(pn_i, v_i n_i)$.
2. We assume that $v_i \sim gamma(q, \gamma)$.

Under these conditions, Fader, Hardie, and Lee (2005b) show the following:

$$\mathrm{E}\left(T_{ij}|p,\, q,\, \gamma,\, \bar{t}_i,\, n_i\right) = \frac{p\left(\gamma + n_i\bar{t}_i\right)}{pn_i + q - 1} = \left(\frac{q-1}{pn_i + q - 1}\right)\frac{p\gamma}{q-1} + \left(\frac{pn_i}{pn_i + q - 1}\right)\bar{t}_i, \tag{2}$$

which provides our desired expression for μ. Estimation of the parameters p, q, and γ will be discussed subsequently. Note that as n_i increases, more weight is placed on the actual observed average IPT $\bar{t}_i$ while less weight is placed on the population mean. Hence, we use the population mean as our expected IPT for customer i when we observe a few transactions from a customer. The more transactions we observe, the more importance is given to the customer-specific average IPT $\bar{t}_i$. Section 5.1 illustrates how this feature leads to individual and adaptive boundaries for our control charts. We then derive a closed-form expression for $\mathrm{Var}(T_{ij}|\bar{T}_i = \bar{t}_i, n_i)$, the full derivation of which is provided in Web Appendix W1:

$$\mathrm{Var}\left(T_{ij}|p, q,\, \gamma,\, \bar{t}_i,\, n_i\right) = \frac{p(p + pn_i + q - 1)\left(\gamma + n_i\bar{t}_i\right)^2}{(pn_i + q - 2)(pn_i + q - 1)^2}. \tag{3}$$

The square root of this expression provides our estimate for σ. To obtain maximum-likelihood estimates of the parameters p, q, and γ, the marginal distribution of $\bar{t}_i$ is required. Fader, Hardie, and Lee (2005b) derive that this marginal distribution is equal to

$$f\left(\bar{T}_i = \bar{t}_i|p, q, \gamma,\, n_i\right) = \frac{\Gamma(pn_i + q)}{\Gamma(pn_i)\Gamma(q)} \frac{\gamma^q \bar{t}_i^{\,pn_i - 1} n_i^{pn_i}}{\left(\gamma + \bar{t}_i n_i\right)^{pn_i + q}} \tag{4}$$

Where Γ is the gamma function. The likelihood to optimize for M customers given their observed number of transactions n_i and average IPT $\bar{t}_i$ is thus given as follows[8]:

$$L\left(p,\, q,\, \gamma|n_i, \bar{t}_i\right) = \prod_{i=1}^{M} f\left(\bar{T}_i = \bar{t}_i|p, q, \gamma,\, n_i\right) \tag{5}$$

[7] We focus on repeat transactions here; hence, T_{ij} denotes the time between the first and second purchases. Customers with only one purchase are excluded from the analysis, as it is not certain that they will make another purchase.

[8] Occasionally updating the model parameters using a new sample of customers more representative of the current customer base may be required as new customers come in and older customers leave the firm.

Based on the parameter estimates obtained by maximizing Equation (5), we compute our quantities of interest in Equations (2) and (3) above, $\mathrm{E}\left(T_{ij}|p,q,\gamma,\bar{t}_i,n_i\right)$ and $\mathrm{Var}\left(T_{ij}|p,q,\ \gamma,\bar{t}_i,n_i\right)$.

3.5. Estimating c for the gamma–gamma model

Having obtained estimates for μ and σ in Equation (1), we now need to obtain the width of the control chart boundaries c. In general, this width can be set to a fixed number, such as 3 (e.g., Shewhart, 1931), but can also be selected based on a sample of observations known to be in-control, combined with a suitable false alarm rate (e.g., P(false alarm) = 0.0027; Montgomery, 2009). As the number 3 is based on the normal distribution theory, and we are dealing with non-normal data in our setting, an in-control sample will be used to determine the value of c. However, a new challenge arises, as the notion of in-control is difficult to define in our case. Based on historical transaction data (e.g., Section 4.1.2), we do not know which customers (or observations from these customers) are behaving as "normal" and which are not. Furthermore, customers vary in transaction frequency, and we cannot reliably determine a false alarm rate due to the short time horizons for several customers. Using only customers with long time horizons biases our estimates, as doing such would ignore a substantial portion of customers.

We develop a simulation approach whereby we generate our own in-control observations based on the distribution of the observed transaction behaviors of customers in the data and determine the value for c based on this simulated sample. For a given set of simulated customers, a series of transaction occasions with IPTs that align with the IPTs observed in the historical data must be generated. To that end, the number of transactions can be generated according to the negative binomial model (NBD), following the Pareto/NBD and BG/NBD models (Fader, Hardie, & Lee, 2005a; Schmittlein et al., 1987).[9]

We obtain the parameters of the NBD model by fitting it to the customer transaction data. Subsequently, this estimated model is used to simulate new, long (10 years) transaction trajectories for 5,000 simulated customers with IPTs that match the behavior of customers in the data. We calibrate the control charts on these simulated transaction trajectories and select c using a grid search, such that most of these observations are within the bounds of the control chart (i.e., are in-control). We follow the existing theory by requiring that c is chosen, such that 99.73% of the observations are within the bounds of the control chart (Shewhart, 1931).

3.6. Model extension: Mixture gamma–gamma model

One assumption underlying our model is that the customer-specific "true" IPT is stable over time. However, research (e.g., Fader, Hardie, & Huang, 2004; Schweidel & Fader, 2009) indicates that customers may show evolving states of transaction or even "hot" and "cold" transaction spells. To accommodate such behavior, we propose a mixture variant of the gamma–gamma model that allows IPT to emerge from a mixture gamma distribution, which formalizes the idea of various IPT states for a given customer[10]. We formulate this variant as follows:

1. We assume that $T_{ij}\sum_{k=1}^{K}w_kZ_{ijk}$ with $Z_{ijk}\sim gamma(p_k,v_{ik})$ and $\sum_{k=1}^{K}w_k=1$. This implies that $E(T_{ij}|p_k,v_{ik})=\xi_i=\sum_{k=1}^{K}w_k\frac{pk}{v_ik}$
2. We assume that $v_{ik}\sim gamma(q_k,\gamma_k)$

[11]In this specification, each observed time may arise from one of the k states, allowing a more flexible IPT distribution. The relevant control chart components become:

$$\mathrm{E}\left(T_{ij}|p,\ q,\ \gamma,\bar{t}_i,\ n_i\right)=\sum_{k=1}^{K}w_k\frac{p_k\left(\gamma_k+n_i\bar{t}_i\right)}{p_kn_i+q_k-1}=\mu \tag{6}$$

$$\mathrm{Var}\left(T_{ij}|p,q,\ \gamma,\ \bar{t}_i,\ n_i\right)=\sum_{k=1}^{K}w_k\left[\frac{p_k(p_k+p_kn_i+q_k-1)\left(\gamma_k+n_i\bar{t}_i\right)^2}{(p_kn_i+q_k-2)(p_kn_i+q_k-1)^2}+\left(\frac{p_k\left(\gamma_k+n_i\bar{t}_i\right)}{p_kn_i+q_k-1}\right)^2\right]-\mu^2. \tag{7}$$

The parameters for this model can be estimated using an expectation–maximization algorithm (Dempster, Laird, & Rubin, 1977) by augmenting Equation (5) with latent class membership variables z_{ik}. We include this model as an additional benchmark in our offline policy analysis (see Section 5.3) because the original data collection was run using the one-segment model.

[9] We choose the NBD model here because it aligns with the distribution of our data. Fig. W3C in Web Appendix 3 shows the alignment between the actual and simulated data. The simulation distribution should be adapted accordingly when the original transaction data have a different transaction distribution.

[10] We thank an anonymous reviewer for suggesting this idea.

[11] We also estimate a model with fixed heterogeneity (i.e., $q_k=q,v_{ik}=v_i$), but this model performs strictly worse in terms of AIC and BIC.

4. Data and empirical setting

We use two empirical datasets and several simulated datasets for our analyses. The first empirical dataset is a historical dataset, which we use to calibrate the control chart using real-life data. The second empirical dataset comes from a field test wherein we assessed the performance of the control-chart approach compared with the current business rule applied by the firm, as well as a control group.

4.1. Real-life setting and empirical datasets

4.1.1. Background of the focal firm

The European online retailer that provided the datasets for this study offers printed greeting card services to customers. Optionally, the customers can decide to send their greeting card with an accompanying larger gift. Similar to many firms, the company is concerned with stimulating customers with reduced purchasing activity levels. Currently, the retailer tries to prevent customers from becoming inactive by sending a so-called reactivation mailing after a customer has not transacted for two months (eight weeks).[12] Such an e-mail primarily serves as a reminder to the customer and offers free postage on the next order, aiming to motivate the customer to purchase from the retailer again.

4.1.2. Historical dataset

We analyze the historical behavior of the 16,790 customers involved in the field test before starting the test. Beyond the transaction data (and variables derived thereof), limited demographic information is also available. Table 2 provides the summary statistics for this dataset. This dataset is used to illustrate the estimation of the control charts in Section 5.1, providing insights in the salient characteristics of the control charts.

4.1.3. Field test dataset

We perform a randomized field test comparing our proposed approach (the model group) to the current status quo at the retailer (the business rule group, sending an e-mail after two months) and a control condition in which no e-mails are sent. Comparisons of the model and business rule groups to the control condition allow us to investigate the relative effect of reactivation targeting customer activity. Fig. 2 presents a graphical overview of the field test and its main outcomes.

The field test is run for two months, from October 1 to November 30, 2016. We choose this period due to the absence of any major holidays (e.g., Valentine's Day, Christmas) or events (e.g., high school graduation ceremonies) that are strongly associated with products from the focal firm and could interfere with the field test. Thus, these months reflect the normal business months for the focal firm[13].

The firm selects 16,790 customers from its database and assigns them randomly to three experimental groups (see Fig. 2). To check whether the randomization is successful, we compare the three groups' key variables in Table 3 and find no significant differences between the groups on average. In Web Appendix W2, we also provide insights into the distribution of these variables across groups, where a few differences are significant for purchase volume, net sales, and frequency. Overall, we conclude that the groups are comparable on most dimensions before the field test.

Customers selected for the field test are randomly drawn from the group of customers that 1) transacted at least once in the previous year and 2) had at least two transactions during their term at the firm. The firm considers the first criterion a good proxy to separate potentially active customers from inactive customers, while the second criterion ensures that only customers who repurchase are included in the field test. We further restrict the eligible group to customers whose final transaction occurred at most two months before the start of the field test and whose control-chart-predicted transaction time fell within the two months of the field test (i.e., we focus on customers at risk of turning inactive). This way, when customers are assigned to either the business rule or the model group, they would actually be approached during the field test.

Given the promising initial results of the analysis of the historical data (see Section 5.1), substantially more customers are allocated to the model group, as the firm is convinced of the performance improvements of our approach and wants to limit the usage of their (potentially poor) approach to reactivation. This is not uncommon in field experiments that are conducted in cooperation with business partners. For example, Ascarza, Iyengar, and Schleicher (2016) assign almost 85% of their customers to the treatment condition. All customers included in the field test are approached only once during the test. Customers do not receive any other targeted marketing actions, including reactivation e-mails, from the firm six months before or during the field test to avoid any confounding effects.

Our field test dataset thus consists of the following information: the assigned group of a customer; whether the customer is active during the field test; when active, the net spending of that customer; and for the model and business rule groups, in which week (1–9) the customer received a reactivation mail.

[12] Discussions with firm managers reveal that two months is chosen because this duration seems "good" and is not based on actual transaction data. Hence, it is independent of the number and timing of transactions, thereby not affecting the observed IPTs for individual customers.

[13] We validate this assertion by applying time-series outlier detection using LOESS seasonal and trend decomposition on aggregate daily sales for the years 2012–2015. There are no outliers during October and November; however, the outliers are concentrated in February (Valentine's) and December (Christmas).

Table 2
Characteristics of the historical dataset.

	Mean	Standard deviation	Minimum	Maximum
Interpurchase time (weeks)[a]	3.57	4.73	0	100
Cumulative purchase volume[a]	3.74	3.29	0.13	100
Cumulative net sales[a]	9.73	6.43	0	100
Relationship length (years)	4.57	2.03	0.07	7.16
Recency (weeks)	6.51	1.18	0.29	8
Frequency	42.74	29.10	4	423
Gender (male)	0.09	0.83	0	1
Greetings-only customer	0.47	0.50	0	1
Greetings-and-gifts customer	0.53	0.50	0	1
Gifts-only customer	0.001	0.008	0	1

[a] This number has been transformed into an index for confidentiality reasons. The index was set to 0 at the 0-point and to 100 at the maximum value.
[b] Net sales are corrected for costs incurred by the firm (e.g., discounts).

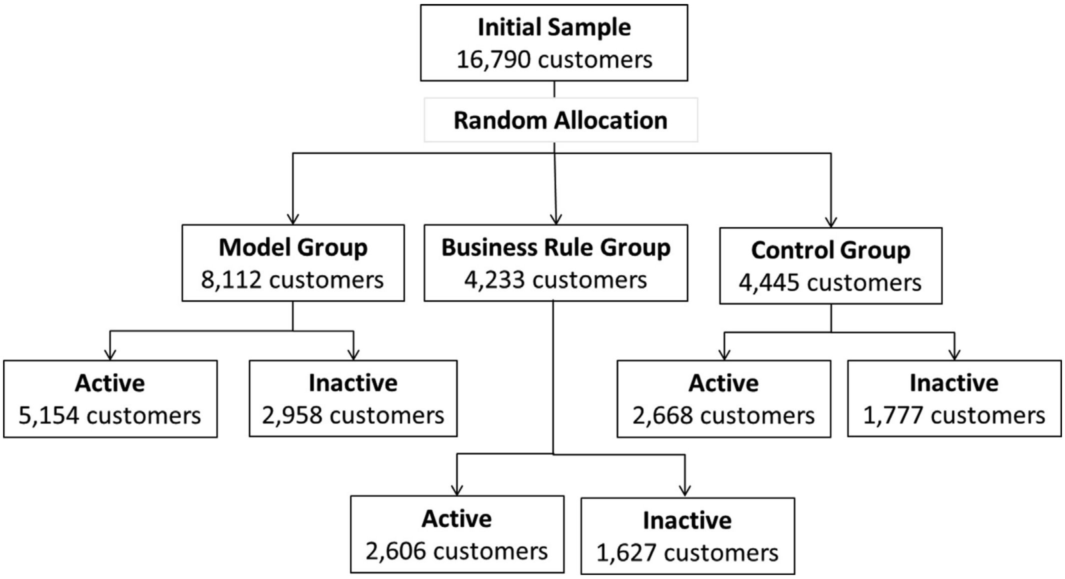

Fig. 2. Overview of the groups in the field test.

Table 3
Pre-test descriptive statistics for the control, model, and business rule group based on historical data.

	Control	Model	Business Rule	Difference p-value
Cumulative purchase volume (index)	3.61	3.88	3.63	0.768
Cumulative net sales (index)	11.51	12.43	11.68	0.379
Relationship length (years)	4.51	4.61	4.55	0.375
Interpurchase time (weeks, index)	11.95	11.56	11.90	0.558
Recency	6.52	6.54	6.43	0.344
Frequency	41.23	42.22	41.50	0.673
N	4,445	8,112	4,233	

Notes: Appendix W2 provides the distribution for each variable.

4.2. Simulated datasets

We use a simulation study to shed light on the performance of our approach under different conditions. We generate 640 synthetic "worlds" consisting of different parameter combinations of the underlying data-generating process, giving rise to 1,600,000 unique simulated customers. Given the importance of variation in transaction timing, we adapt the simulation design of Platzer and Reutterer (2016), who built on the design of Fader et al. (2005a). The details of our simulation approach are provided in Web Appendix W4.

In our simulation, we estimate a series of models on simulated calibration data and subsequently use these models to make holdout predictions in terms of customer activity. We discuss our benchmark models after the analysis of the field test data. Following Korkmaz et al. (2013), we generate individual-level timing predictions for these benchmark models and compare them to the actual holdout activity time. For each benchmark model, we compute the relative lift in mean average error (MAE) as MAE_lift = 1 – MAE_{cchart}/MAE_{bench}, where MAE is defined as MAE = $\sum_{i=1}^{N}|x_i^* - x_i|$ with x_i^* the first predicted holdout activity time and x_i as the actual activity time. Higher lift values indicate better performance of the control chart over the specific benchmark.

5. Results

5.1. Illustrating control-chart estimation and salient features

We use the historical dataset to calibrate the control charts to illustrate their estimation and usage. We estimate the control charts for the 8,112 customers in the model group of the field test. First, the parameters for the gamma–gamma control chart are obtained by optimizing Equation (5). These estimates correspond to $\widehat{p}$ = 15.73, $\widehat{q}$ = 1.62, and $\widehat{\gamma}$ = 0.57. Afterward, the boundaries for the control chart are determined using the procedure outlined in Section 3.5, obtaining the estimate $\widehat{c}$ = 4.4. These four parameters specify the control chart, which we can compute for every customer by plugging Equations (2) and (3) into Equation (1).

The boundaries in our case are slightly wider compared with the case of the normal distribution ($\widehat{c}$ = 3; Shewhart, 1931). Shewhart (1931) assumes that the sequence of observations is independent, which is not the case in our situation; thus, we need wider control limits (Wieringa, 1999). Fig. W3C in Web Appendix 3 shows that the mass of the simulated data aligns with the actual data, but as intended, the simulated data contain more customers with longer transaction horizons. This condition aligns with the idea that some customers still need to settle into a regular transaction cycle, for example, due to being in a trial phase (Schweidel & Fader, 2009).

Fig. 3 shows two control charts: one for a customer with below-average IPT (customer A) and one for a customer with above-average IPT (customer B). The former illustrates what happens for a customer who has short transaction cycles, while the latter illustrates what happens for a customer with long transaction cycles.

Both plots illustrate the salient characteristics of the control chart. First, the time since the last transaction increases each week until a new transaction occurs and then drops to zero. Second, both plots show the estimated average time between transactions (Equation (2) as a solid line that evolves when new information arrives. Third, and most importantly, the upper bound for the control chart is visible. Two important characteristics of the boundaries are that a) they differ between customers and b) they change over time for individual customers. Both characteristics emerge from Equations (2) and (3). The differences between customers arise due to differing IPTs ($\bar{t}_i$), while changes over time within customers arise due to updates when new IPTs (T_{ij}) are observed, providing additional information on a customer's normal transaction pattern. Consistent with Equation (2), the starting boundaries are wider due to reliance on customer-base level estimates and then adapt to more customer-specific boundaries when transactions occur. This characteristic also allows the control chart to adapt to time-related events that affect a customer's normal transaction pattern (e.g., seasonality, special events) by shifting the boundaries once a new transaction has occurred. This is, for example, apparent for customer A around week 25, where a permanent upward boundary shift potentially caused by a special event is observed. Furthermore, this customer exhibits some seasonality in its transaction pattern, which we consider a common cause variation (see Section 2.3) and, thus, should not be a reason for reactivation. The control-chart boundaries adapt to this pattern by including these observations inside the boundaries.

The boundaries of the plots also differ. Although the time between transactions for customer B (lower panel) never crosses the upper boundary of the control chart, the chart for customer A crosses the boundary several times. Thus, while we would never attempt reactivation for customer B, we would do so for customer A, specifically at the point in time at which the boundary is crossed. It is at that point that we are unsure whether the customer will transact again, and reactivation may stimulate a transaction. Intuitively, this difference between both customers makes sense. Given the long period between customer B's transactions, we can still reasonably expect a transaction to occur in the future. Reactivation would likely not be effective, given this customer's normal transaction pattern.

By contrast, given the short transaction cycles of customer A, a longer time between transactions (e.g., around week 75) can be indicative of the customer becoming inactive, and reactivation can be warranted.

5.2. Field test results

5.2.1. Model-free analysis

Our outcome variable of interest is a binary activity indicator: whether or not the customer transacted during the field test. This is in line with our and the firm's objective to reactivate customers who did not transact recently and, thus, can be at risk of not transacting again in the future. The randomized nature of our field test enables direct activity comparisons of customers across the three conditions to gain model-free insights into the field test results (see Fig. 2).

On average, model-based targeting is more effective than the business rule targeting method: 63.5% of the customers in the model group transacted compared with 61.6% of the customers in the business rule group and 60.0% of the customers in

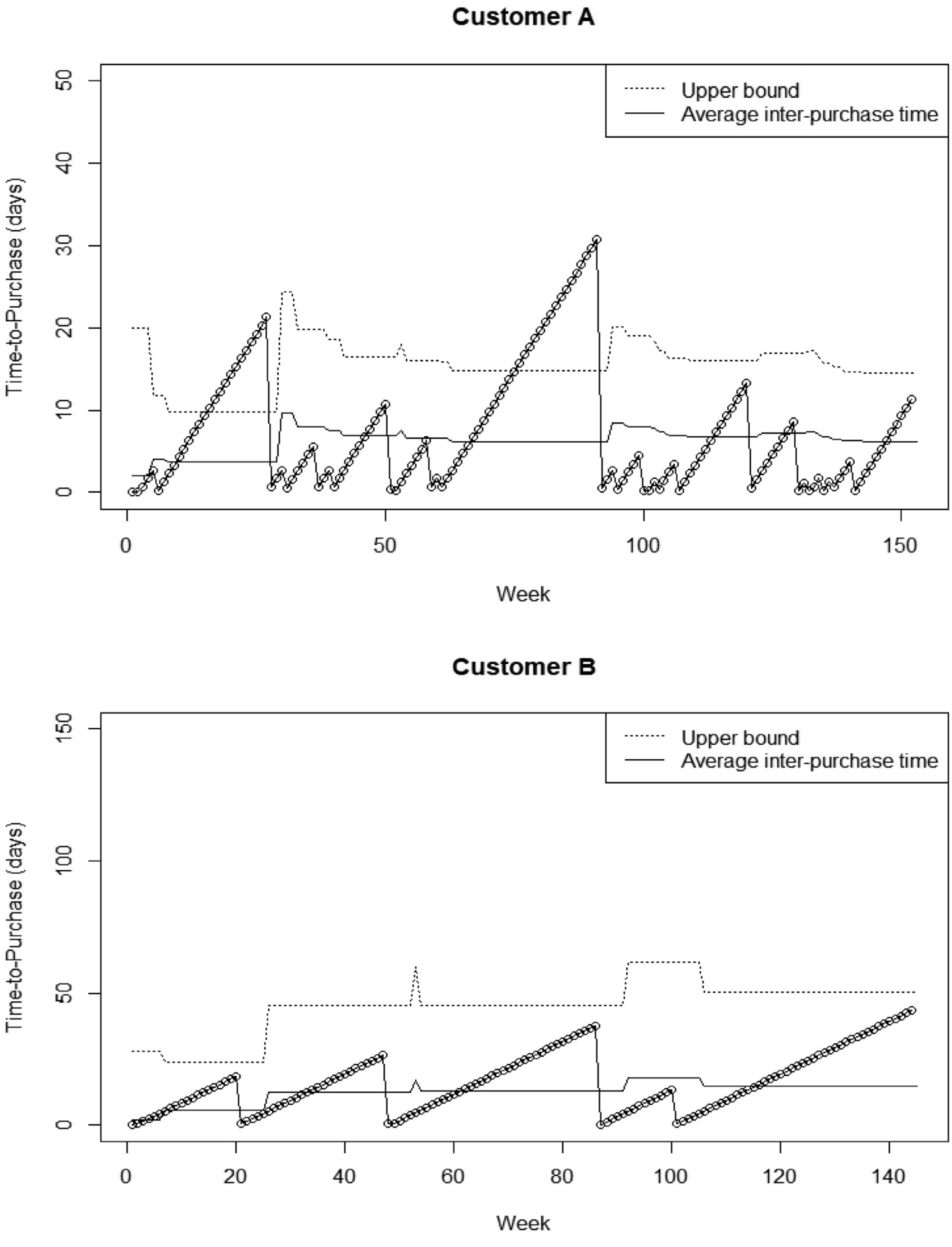

Fig. 3. Control charts for two randomly selected customers A and B.

the control group. A three-way chi-square test indicates significant differences between groups ($\chi^2(2) = 15.77$, $p < 0.01$). A follow-up pairwise proportion test (corrected for multiple comparisons using Benjamini and Hochberg's (1995) method) confirms that the difference between the model and the business rule groups is significant ($p = 0.049$), and the difference between the model and control groups is significant ($p < 0.01$). No significant difference has been observed between the business rule and the control groups ($p = 0.148$). Therefore, on average, the model-based targeting approach outperforms the business rule approach while also increasing the total number of active customers compared with not taking any action.

5.2.2. Model-based analysis

While these results provide initial evidence of the effectiveness of our reactivation approach, they do not control for potential confounding factors that influence activity, nor do they consider customer heterogeneity. Therefore, we also estimate a formal model relating activity to a set of control factors. We control for variables typically used in the CRM literature, namely, relationship length, IPT and RFM variables[14], and gender (Blattberg, Kim, & Neslin, 2008). Furthermore, we explore heterogeneity in the treatment effects by adding the interactions between the dummies indicating the groups and the control factors to our model. We estimate the following linear probability model:

$$\text{Activity}_i = \beta_0 + \beta^m TM_i + \beta^b TB_i + \beta^c X_i^c + \varepsilon_i, \tag{6}$$

where Activity_i is a binary variable that indicates whether a customer transacted during the experiment, TM_i and TB_i are binary variables that indicate the model and business rule groups, respectively, and X_i^c is a vector containing the control variables (relationship length, IPT, RFM, and gender). The vector β contains the intercept β_0, the average treatment effects for the model and business rule groups β^m and β^b, and the main effects for the control variables β^c. Finally, the error term ε_i is assumed to follow a normal distribution. Continuous variables are mean-centered to represent the effects for the "average" customer.

Table 4 presents the results for the linear probability models. The results in the first column confirm that, compared with the control group, activity is significantly higher for the model group but not for the business rule group. Using the business rule group as the baseline for our treatment variables (Column 3) reveals that the difference between the model and the business rule groups is significant ($b = .024, p = .007$). The second column shows the robustness of these effects when controlling for pre-test factors that potentially influence activity. This is reassuring and suggests that our result is not an artifact of a failure of randomization. The difference between the model and business rule groups remains significant ($b = .031, p = .003$, column 4).

Some pre-test factors also influence activity directly. We find that customers who have been with the firm longer and those with a higher transaction frequency have a higher probability of transacting during the field test period. By contrast, customers with a higher average IPT, greetings-only customers, customers who transacted longer ago, and male customers have a lower probability of purchasing during the field test.

Overall, we confirm the findings of the previous section that, on average, the model-based targeting method significantly increases the transaction probability of customers.

5.3. Comparison with existing approaches using offline policy analysis

How well does the control-chart approach outlined above fare in separating active customers from inactive customers when we compare its performance with that of existing approaches identified in the literature? We investigate this by splitting our samples into training and holdout samples (e.g., Fader, Hardie, & Lee, 2005a) and investigating how well different approaches predict the likelihood of activity (i.e., risk) in the holdout sample. However, this ignores the fact that our decision to reactivate a customer is based on the presumption that sending a reactivation mailing generates a positive incremental response (i.e., positive lift), that is, the customer resumes purchasing *due to* the intervention. Thus, as Ascarza (2018) and Lemmens and Gupta (2020) argue, we should compare models on their ability to generate lift, which can be done using (field) experiments.

Ideally, we would run experiments with all possible benchmarks as conditions, but such is often expensive and infeasible. For example, our field test only includes two active conditions (the control chart and the business rule), which means we cannot compare our approach directly with other existing methods, as discussed in Section 2.

To overcome this shortcoming, we follow the aforementioned studies and conduct an offline policy analysis (Li et al., 2012; Hitsch and Misra, 2018; Yoganarasimhan et al., 2022). Offline policy analysis allows for the arbitrary testing of many other targeting policies—in our case, based on different models from prior literature—using data from only one randomized test that did not include these policies as conditions. The key insight of Hitsch and Misra (2018) is that only the usable observations from the randomized test are used (i.e., those observations for which the targeting policy of the randomized test is the same as for the policy we want to test against). Invariably, this leads to a loss of observations. However, Hitsch and Misra (2018) show that we can correct for this loss of observations by weighting the usable observations with an inverse probability weight (e.g., Horvitz & Thompson, 1952; Robins, Rotnitzky, & Zhao, 1994) based on the probability of being usable.

Formally, let $T_i \in \{0, 1\}$ denote whether customer *i* is targeted during the field test or not. Based on the targeting status, we derive the *activity* of customer *i* as follows:

$$\pi_1(T_i) = \begin{cases} Y_i(0) & \text{if } T_i = 0 \\ Y_i(1) & \text{if } T_i = 1 \end{cases}$$

and the associated *net profit* for a margin *m* and targeting cost *c* as

[14] Only information on *cumulative* net sales is available. Hence, we use a categorical variable with the categories greetings-only, gifts-only, and mixed customers for monetary value to avoid collinearity problems with relationship length and frequency. The two latter categories have higher value customers, given the price difference between greetings and gifts.

Table 4
Linear probability model results for customer activity drivers.

	(1) **Activity (main effect)**	(2) **Activity (controls)**	(3) **Activity (main effect, business rule baseline)**	(4) **Activity (controls business rule baseline)**
Intercept	0.600 (0.007) ***	0.613 (0.007) ***	0.615 (0.007) ***	0.635 (0.008) ***
Control group	0 (baseline)	0 (baseline)	−0.015 (0.010)	−0.012 (0.010)
Model group	0.035 (0.009) ***	0.042 (0.010) ***	0.024 (0.003) ***	0.031 (0.010) ***
Business rule group	0.016 (0.010)	0.011 (0.010)	0 (baseline)	0 (baseline)
Gender (male)		−0.023 (0.013)		−0.021 (0.013) *
Relationship length[a]		0.005 (0.002) *		0.012 (0.002) ***
Interpurchase time[a]		−0.003 (0.002) *		−0.018 (0.002) ***
Greetings-only customer		−0.029 (0.007) ***		−0.029 (0.007) ***
Gifts-only customer		−0.607 (0.482) *		−0.643 (0.481)
Recency[a]		−0.004 (0.000) ***		−0.004 (0.000) ***
Frequency[a]		0.001 (0.000) ***		0.000 (0.000) **
N	16,790	16,790		
Adjusted R^2	0.0001	0.016		

Notes: [a] All continuous variables are mean-centered, so the main effects refer to the average customer. Standard errors in parentheses. *** p < 0.01; ** p < 0.05; * p < 0.10.

$$\pi_2(T_i) = \begin{cases} mY_i(0) & \text{if } T_i = 0 \\ mY_i(1) - c & \text{if } T_i = 1 \end{cases}$$

We now introduce a targeting policy $p : X_i \to \{0, 1\}$, which indicates whether a customer with characteristics X_i should be targeted, $p(X_i) = 1$, or not, $p(X_i) = 0$. We then observe the total activity ($j = 1$) or net profit ($j = 2$) for all N customers in the field test as follows:

$$\hat{I}_j(p) = \sum_{i=1}^{N} \left((1 - p(X_i)) * \pi_j(0) + p(X_i) * \pi_j(1)\right). \tag{8}$$

Hitsch and Misra (2018) show that for targeting policies p not being part of the field test, we can compute Π using those observations for which $T_i = p(X_i)$. For several observations, however, it will hold that $T_i \neq p(X_i)$. We can correct Equation (5) for this using inverse probability weights based on the probability that a customer is targeted in the field test $0 < \varepsilon < 1$:

$$\Pi_j(p) = \sum_{i=1}^{N} \left(\frac{1 - T_i}{1 - \varepsilon}(1 - p(X_i)) * \pi_j(0) + \frac{T_i}{\varepsilon} p(X_i) * \pi_j\ (1)\right) \tag{9}$$

Using Equation 9, we compute the activity (net profit) for a variety of targeting policies p, estimating ε using a logistic regression model and the customer covariates available. For the targeting cost c, we consider four values: 0.01, 0.33, 0.605 and 0.938. The first value equals the cost of a regular reminder e-mail without incentive[15], the second value equals that of a typical 10% discount given by the firm, the third value equals the cost of free postage, and the fourth value combines a 10% discount with free postage. We select p based on models previously identified in Section 2:

1. Pareto/NBD (Schmittlein et al., 1987)
2. BG/NDB (Fader, Hardie, & Lee, 2005a)
3. Pareto/GGG (Platzer & Reutterer, 2016)
4. Pareto/NBD with reactivation (Ma, Tan, & Shu, 2015)
5. MBG/CNBD-k (Reutterer, Platzer, & Schröder, 2021)
6. HMM (Schwartz, Bradlow, & Fader, 2014)
7. Generalized gamma model (Allenby, Leone, & Jen, 1999)
8. Survival model, i.e., heterogeneous discrete-time proportional hazard model with expo-power baseline hazard (Seetharaman & Chintagunta, 2003)
9. Uplift models (Guelman, Guillén, & Pérez-Marín, 2015; Ascarza, 2018)
10. Causal random forest (Wager & Athey, 2018)
11. Causal survival model (Cui et al., 2020)

Broadly, we can classify these models as models built directly upon the time dimension (Models 7, 8, and 11), models that provide individual-level parameters or predictions (Models 3, 6, 9, and 11), and traditional stochastic models that estimate aggregate-level parameters (Models 1, 2, 4, and 5). Our approach should foremost be compared with the first set of models,

[15] https://www.emailvendorselection.com/cost-per-mille-cpm/.

given their focus on predicting transaction *time* rather than transaction *propensity*. One exception is the HMM from the second set, which can also directly estimate transaction time (but does not specifically build upon the time dimension).

The causal random forest (Model 10) and causal survival model (Model 11) are newer variants of the uplift model. We include these models because while the uplift model (Model 9) maximizes the lift impact of targeting (i.e., whom to target), 1) it is sensitive to non-random treatment assignment (Zhang, Li, & Liu, 2021) and 2) does not provide an answer to the question of when to target. Causal random forests address the first issue by including propensity scores, correcting for potential non-random treatment assignments based on observables. The survival model also addresses the second issue by including a time-to-target estimate (i.e., the remaining survival time) while still targeting based on the treatment effect.

A suitable targeting policy p requires transforming model predictions to a binary classification signaling whether a customer would be targeted during the field test period or not. Models 6, 7, 8, and 11 provide direct estimates of the time at which a customer should be targeted, and $p = 1$ if this time falls within the field test period. For Models 1–5, we rely on the method developed by Korkmaz et al. (2013) to derive individual-level transaction timing estimates, which we translate to a policy p by setting $p = 1$ if a transaction timing estimate is within the field test period but transaction did not occur by the predicted time. For Model 9, we rely on uplift random forests (Guelman, Guillén, & Pérez-Marín, 2015) that provide individual-level estimates for two probabilities: the probability of activity in case of 1) targeting and 2) not targeting. Subtracting these probabilities yields the incremental effect (or lift). The corresponding targeting policy p is then directed at customers for which lift is positive. The policy for Model 10 is derived similarly using models tuned with 10-fold cross-validation for the hyperparameters. Finally, we add the control chart, mixture control chart, and business rule to our comparison.

For our offline policy evaluation, we use data from the business rule and control groups only. For the latter group, we only include customers who are inactive for eight weeks to ensure comparability between groups. We exclude the control chart group as it was determined by our approach, and we want to avoid biasing our findings. We use a bootstrap approach with 100 training and holdout samples, estimating our models on 50% of customers and then using the remaining 50% to generate the predictions based on Equation 9. Models 1, 2, 4, and 6 include the treatment group variables, while Models 7–11 include the pre-test covariate information (including a treatment group indicator) to ensure the comparability of the models. We report in Table 5 the results for the most directly comparable models (the time dimension models) and the best-performing models from the other group. Appendix 5 provides the full results for all models.

Our control chart approach outperforms almost all other time-dimension models in terms of the number of active customers (3,331 or 65.5%) and net profit (€10,957.99). However, two exceptions occur. First, while the control chart identifies more active customers, it does not outperform the best-fitting mixture control chart (with k = 2 segments) in terms of profit. Second, the control chart shows equivalent performance to the causal survival model on both activity (3,229 or 63.5%, $p = 0.868$) and profit (€10,726.02; €9,198.95; $p > 0.10$), except for profits in the highest-cost scenarios (€329.19, $p < 0.00$, (€382.73, $p < 0.00$).

With respect to other model types (particularly the latent attrition models Pareto-NBD and HMM), we find that the control chart performs better in terms of activity but only in terms of profit when costs are higher. Only one model strictly outperforms the control chart model in terms of active customers, namely, the uplift model, with a significantly higher number of active customers (4,576 or 90.1%). At the same time, while net profit is significantly higher in the lowest-cost scenario (€15,826.73), this profit difference disappears in when reactivation costs increase, and the control chart even generates significantly higher profits in the highest cost scenario (€9,828.10 versus €7,718.43.10, $p < 0.00$). This result confirms earlier findings (Ascarza, 2018; Lemmens & Gupta, 2020) that targeting lift instead of risk (as the control-chart approach does) can yield better performance in terms of activity (but not profit per se). Notably, however, the uplift approach only provides insight on *whom* to target and, thus, does not directly meet the objective we set for our approach (determining *when* to target). If we compare the control-chart approach with the causal survival model, which is the time-dimension model closest to the uplift model, the difference between both approaches disappears. Therefore, we can conclude that the control-chart approach approximates lift-based targeting while also indicating when a customer should be targeted.

5.4. Conditions influencing control-chart performance

The previous sections highlight that the control chart approach is a competitive alternative to most approaches when it comes to predicting the timing of reactivation interventions. However, under what conditions does our approach perform better or worse? We shed light on this question using the results from the simulated datasets. For our simulation, we benchmark against one model out of each category in Table 5. We select causal survival, HMM, and Ma et al.'s (2015) Pareto-NBD model. We do not include the uplift and causal random forests given the binary (not time-scale) nature of their predictions.

Figs. 4 and 5 summarize our findings. Similar to Reutterer, Platzer, and Schröder (2021), we report our results in terms of CART models.

We explain the lift metric for each benchmark model in terms of the shape and rate simulation design parameters for the regularity parameter $k \sim Gamma(t, \gamma)$, the transaction rate $\lambda \sim Gamma(r, \alpha)$, and the dropout rate $\mu \sim Gamma(s, \beta)$, as well as the number of weeks in the training and test set, respectively. The Pareto-NBD model is discussed in Web Appendix W6 for space reasons.

Table 5
Best-performing benchmark models in offline policy analysis.

	Holdout # active customers (sd)	Holdout % activity (sd)	Holdout net profit (sd)[b] (c = 0.01)	Holdout net profit (sd)[b] (c = 0.33)	Holdout net profit (sd)[b] (c = 0.605)	Holdout net profit (sd)[b] (c = 0.938)
Models accounting for time dimension						
Control chart	**3331 (88.5)**	**0.655 (0.116)**	**10,957.99 (690.5)**	**10,859.19 (1404.6)**	**10,392.84 (1167.64)**	**9,828.10 (893.8)**
Control chart (mixture, k = 2)[a]	2995 (145.7)*	0.590 (0.285)	10,091.48 (493.8)	9,304.41 (394.4)	8,628.92 (309.6)	7,949.21 (285.8)
Two-month hiatus	2582 (37.2)**	0.508 (0.007)	8,812.38 (149.1)*	7,669.16 (148.3)*	6,686.70 (147.9)**	5,497 (147.8)**
Causal survival	3229 (149.5)	0.635 (0.029)	10,726.02 (545.5)	9,198.95 (465.5)	329.19 (242.5)**	382.73 (18.34)**
Survival model	2584 (55.9)**	0.509 (0.011)	8,809.51 (222.5)*	5,075.53 (144.1)**	4,356.15 (128.9)**	3,870.07 (177.9)**
Generalized Gamma	2472 (35.6)**	0.487 (0.007)	8,415.76 (152.8)*	7,156.57 (149.1)**	6,070.53 (146.5)**	5,414.56 (450.1)**
Individual level prediction models						
Uplift Random Forest	4576 (319.2)**	0.901 (0.063)	15,826.73 (1047)*	13,593.18 (849.5)†	11,658.92 (698.9)	7,718.43 (314.2)*
HMM	2897 (41.64)*	0.570 (0.008)	9,955.87 (179.9)†	8,450.42 (175.7)†	7,119.45 (164.6)**	6,477.29 (673.0)**
Aggregate level prediction models						
Pareto-NBD (Ma et al.)	2992 (229.6)**	0.589 (0.045)	10,935.10 (1018.1)	8,151.67 (763.5)**	7,892.89 (574.1)**	7,476.41 (468.7)*

Notes: The holdout active customers (activity percentage) column reports the mean number (percentage) of predicted active holdout customers across bootstrap iterations under a specific policy. The holdout net profit column reports the associated net profit of these customers for three different cost parameters: 0.01, 0.33 and 0.605. P-values computed from the bootstrap difference between the benchmark model and the control chart with respect to either number of active customers or net profit.
† $p < 0.10$; * $p < 0.05$; ** $p < 0.01$.
[a] This was the best fitting mixture control chart based on AIC and BIC.
[b] Profit values are scaled by a constant to maintain confidentiality.

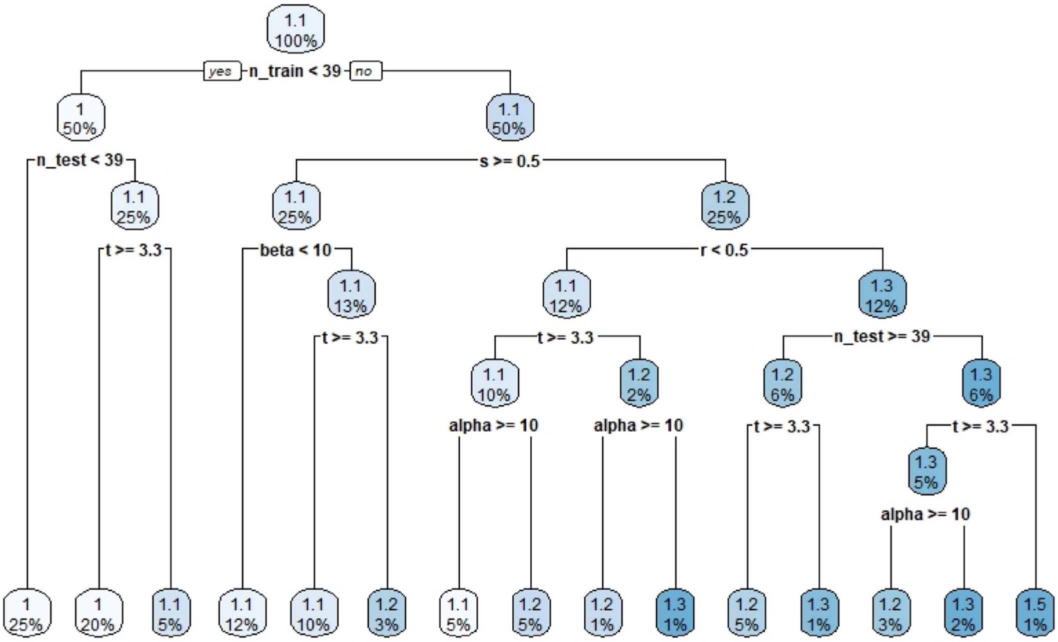

Fig. 4. Regression tree for lift versus the causal survival model.

Benchmarking against the causal survival model. Fig. 4 summarizes our results for the causal survival model. The control chart performs very similarly to the causal survival model for 45% of the cases (lift = 1), mainly characterized by short (26 weeks) training and test periods. Performance improvements (lift > 1) arise only when the control chart has sufficient (>26) weeks to calibrate. In these cases, performance improvements arise when the dropout rate shape is < 0.5 and the trans-

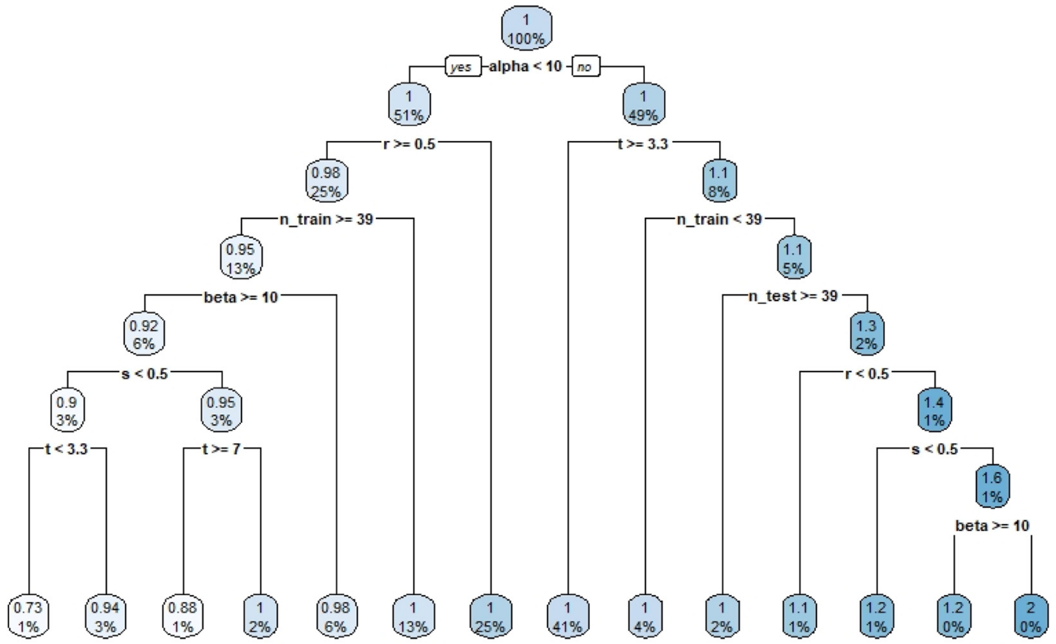

Fig. 5. Regression tree for lift versus the HMM.

action rate shape r > 0.5, as indicated by the terminal nodes on the right side of Fig. 4. The coefficient of variation for the gamma distribution is $1/\sqrt{shape}$. Thus, higher shape parameters indicate more homogeneity in the distribution. The strongest improvements thereby occur for more homogeneous transaction distributions and for more heterogeneous dropout distributions. As a result, relative to the causal survival model, the control chart can better deal with heterogeneous dropout, which aligns with our findings in Section 5.1 regarding control chart adaptability. However, transaction distributions should be more homogeneous, and there should be a sufficient number of weeks prior to when predictions are made to calibrate the control chart. Otherwise, the performance between models is equivalent.

Benchmarking against the HMM. Fig. 5 summarizes our benchmark results for the HMM. Improvements in the control chart relative to the HMM are driven primarily by lower values of the regularity distribution shape parameter ($t < 3.3$, right side of Fig. 5). This result indicates a relatively better improvement for heterogeneous transaction regularity distributions. Although the control chart shows similar performance to the HMM for shorter train and test time horizons (<39 weeks), the performance of the control chart improves notably with longer time horizons. The reduced performance of the HMM in this setting aligns with the findings of Netzer, Srinivasan, and Lattin (2008), who report that HMM predictions suffer in longer and heterogeneous samples because the model struggles to leverage heterogeneity in holdout predictions.

Overall, the results of the simulations (including the Pareto/NBD reported in the Web Appendix) show that the control chart can systematically outperform the benchmarks, mostly in situations where heterogeneity in the transaction and/or dropout processes play a strong role. We attribute this to the model's adaptability to individual customer's transaction trajectories. When transaction and/or dropout distributions are more homogeneous, benchmark models should be preferred over the control chart. Finally, for shorter calibration and prediction periods, the control chart performs mostly equivalent to the benchmarks; its performance improves with longer time horizons.

6. Discussion

6.1. Extending and corroborating extant research

Customer reactivation refers to the proactive communication with customers who reduced their transaction frequency temporarily or stopped their transactions entirely (Blömeke et al., 2010). To aid marketing practitioners struggling with implementing customer reactivation, we introduce a gamma-gamma control chart model for customer reactivation timing, combining insights from customer management with those from statistical quality control theory. This model addresses the shortcoming of existing models that have poor performance when predicting individual customer activity (Wübben & Von Wangenheim, 2008).

A field experiment establishes the efficacy of our approach. Compared to a control group, customer activity increased by 3.5 percentage points. Relative to the current business rule, this increase was 1.9 percentage points. Hereby, we extend the finding of Drèze and Bonfrer (2008), who show the aggregate level impact of mailing timing, to the individual level. While the average effect is positive, exploratory analyses in Web Appendix W7 uncover treatment effect heterogeneity (Luo, Lu, & Li 2019; Wager & Athey, 2018), indicating that further activity gains can be obtained by targeting specific subgroups.

We also benchmark the performance of the control chart approach to existing models using offline policy evaluation (Hitsch & Misra, 2018) and simulations. Offline policy analysis shows that existing stochastic models for customer base analysis (Table 5) struggle with predicting individual activity, and have lower profits. Although such models provide good tools for customer base management, they are thus less suited for reactivation targeting. Our control-chart approach generates a larger number of active customers compared to most benchmarks, although this does not always increase profits. These observations confirm the finding of Drèze and Bonfrer (2008) that activity and revenue behave asymmetrically under adapted timing. Finally, simulation results show that the control chart performs better than the benchmarks when transaction and/or dropout distributions are more heterogeneous rather than homogeneous, and when time horizons are longer ($\geq$52 weeks).

Our findings corroborate existing research in three ways. First, we confirm earlier studies showing that lift-based methods outperform risk-based methods (e.g., Ascarza, 2018; Lemmens & Gupta, 2020). Indeed, an uplift model generates more active customers than the control chart (Table 5). Importantly however, 1) this only translates into profit increases for low reactivation costs, and 2) it does not inform managers *when* to target a customer, which our control chart does, and better so than its lift-based equivalent (causal survival model). Second, corroborating the findings of Ascarza, Iyengar, and Schleicher (2016), in Web Appendix W8 we show that proactively targeting customers *too early* (rather than just targeting them) reduces transaction probability. Targeting *too late* has no significant impact, in line with Drèze and Bonfrer (2008), who find that long intercommunication times are less problematic than short ones. Third, corroborating Kumar, Bhagwat, and Zhang (2015, p. 52), we find that reactivation incentives matter, as they generate different firm profit outcomes (Table 5).

6.2. Limitations and future research

Within the scope of this paper, some limitations remain. First, our field experiment did not investigate the impact of 1) sending multiple reactivation mailings, and 2) sending reactivation mailings based on expected profit (e.g., Lemmens and Gupta 2020), which future research could investigate. Second, sample selection for the field experiment was restricted to ensure comparability between experimental groups (Section 5.2), limiting the interpretation of our results to a subgroup of customers. Third, our model does not incorporate covariates, but Web Appendix W9 sketches how this could be done. Fourth, customer transaction behavior could endogenously change due to our new targeting approach (Lewis, 2005), although we deem this unlikely in our setting as customers would find it difficult to infer the reactivation targeting approach from the other mailing signals sent by the firm (Meyer and Hutchinson 2016).

Data availability

Example code is available at the corresponding author's personal page. Data used for the paper are confidential otherwise.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is part of the first author's dissertation written at the University of Groningen. The authors would like to thank the anonymous retailer for its cooperation in this study, as well as the area editor and two anonymous reviewers for their comments. Dylan Beijers, Arend-Jan Hengeveld, Aurélie Lemmens, and Thomas Reutterer provided helpful comments on earlier versions of this manuscript. The participants at the 2016 and 2019 EMAC Conferences, 2016 and 2019 Marketing Dynamics Conferences, and 2018 Marketing Science Conference, as well as the seminar participants at Goethe University, KU Leuven-Groningen Research Seminar, Maastricht University, University of Amsterdam, and University of Groningen, also provided valuable feedback on this work. Example code is available at the corresponding author's webpage.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijresmar.2023.05.001.

References

Allenby, G. M., Leone, R. P., & Jen, L. (1999). A dynamic model of purchase timing with application to direct marketing. *Journal of the American Statistical Association, 94*(446), 365–374.
Ascarza, E. (2018). Retention futility: Targeting high-risk customers might be ineffective. *Journal of Marketing Research, 55*, 80–98.

Ascarza, E., & Hardie, B. G. S. (2013). A joint model of usage and churn in contractual settings. *Marketing Science, 32*(4), 570–590.
Ascarza, E., Iyengar, R., & Schleicher, M. (2016). The perils of proactive churn prevention using plan recommendations: Evidence from a field experiment. *Journal of Marketing Research, 53*(1), 46–60.
Ascarza, E., Neslin, S. A., Netzer, O., Anderson, Z., Fader, P. S., Gupta, S., ... Schrift, R. (2018). In pursuit of enhanced customer retention management: Review, key issues, and future directions. *Customer Needs and Solutions, 5*, 65–81.
Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B, 57*, 289–300.
Blattberg, R. C., Kim, B.-D., & Neslin, S. A. (2008). *Database marketing: Analyzing and managing customers*. Springer Science + Business Media.
Blömeke E., Clement, M., & Bijmolt, T. H. A. (2010). *Should they stay or should they go? Reactivation and termination of low-tier customers: Effects on satisfaction, word-of-mouth, and purchases*. SOM Research Reports (Vol. 10008). University of Groningen, SOM Research School.
Bult, J. R., & Wansbeek, T. (1995). Optimal selection for direct mail. *Marketing Science, 14*(4), 378–394.
Colombo, R., & Jiang, W. (1999). A stochastic RFM model. *Journal of Interactive Marketing, 13*, 2–12.
Crespy, C. T., Stearns, J. M., & Krehbiel, T. C. (1995). Integrating analytic methods into marketing research education: Statistical control charts as an example. *Marketing Education Review, 5*(1), 11–23.
Cui, Y., Kosorok, M. R., Sverdrup, E., Wager, S., & Zhu, R. (2020). Estimating heterogeneous treatment effects with right-censored data via causal survival forests. *arXiv preprint arXiv:2001.09887*.
Deming, W. E. (1982). *Out of the crisis*. Cambridge University Press.
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), 1–22.
Does, R. J. M. M., Roes, K. C. B., & Trip, A. (1999). *Statistical process control in industry: Implementing and assuring SPC*. Kluwer Academic.
Drèze, X., & Bonfrer, A. (2008). An empirical investigation of the impact of communication timing on customer equity. *Journal of Interactive Marketing, 22*(1), 36–50.
Ernst, R., Guerrero, J.-L., & Roshwalb, A. (1993). A quality control approach for monitoring inventory stock levels. *Journal of the Operational Research Society, 44*(11), 1115–1127.
Fader, P. S., & Hardie, B. G. S. (2013). *The Gamma–Gamma model of monetary value*. http://brucehardie.com/notes/025.
Fader, P. S., Hardie, B. G., & Huang, C. Y. (2004). A dynamic changepoint model for new product sales forecasting. *Marketing Science, 23*(1), 50–65.
Fader, P. S., Hardie, B. G. S., & Lee, K. L. (2005a). Counting your customers the easy way: An alternative to the Pareto/NDB model. *Marketing Science, 24*(2), 275–284.
Fader, P. S., Hardie, B. G. S., & Lee, K. L. (2005b). RFM and CLV: Using iso-value curves for customer base analysis. *Journal of Marketing Research, 42*, 415–430.
Gönül, F., Kim, B.-D., & Shi, M. Z. (2000). Mailing smarter to catalog customers. *Journal of Interactive Marketing, 14*(2), 2–16.
Gönül, F., & Shi, M. Z. (1998). Optimal mailing of catalogs: A new methodology using estimable structural dynamic programming models. *Management Science, 44*(9), 1249–1262.
Gönül, F., & ter Hofsede, F. (2006). How to compute optimal catalog mailing decisions. *Marketing Science, 25*(1), 65–74.
Guelman, L., Guillén, M., & Pérez-Marín, A. M. (2015). Uplift random forests. *Cybernetics and Systems, 46*(3/4), 230–248.
Hitsch, G. J., & Misra, S. (2018). Heterogeneous treatment effects and optimal targeting policy evaluation. *Available at SSRN 3111957*.
Horvitz, D. G., & Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association, 47*(260), 663–685.
Korkmaz, E., Kuik, R., & Fok, D. (2013). *Counting your customers: When will they buy next? An empirical validation of probabilistic customer base analysis models based on purchase timing*. ERIM Report Series ERS–2013–001–LIS. https://ssrn.com/abstract=2198260.
Kumar, V., Bhagwat, Y., & Zhang, X. (2015). Regaining "lost" customers: The predictive power of first-lifetime behavior, the reason for defection, and the nature of the win-back offer. *Journal of Marketing, 79*(4), 34–55.
Lemmens, A., & Gupta, S. (2020). Managing churn to maximize profits. *Marketing Science, 39*(5), 956–973.
Lewis, M. (2005). Incorporating strategic consumer behavior into customer valuation. *Journal of Marketing, 69*(4), 230–238.
Li, L., Chu, W., Langford, J., Moon, T., & Wang, X. (2012). May). An unbiased offline evaluation of contextual bandit algorithms with generalized linear models. In *Proceedings of the Workshop on On-line Trading of Exploration and Exploitation 2, JMLR Workshop and Conference Proceedings* (pp. 19–36).
Luo, X., Lu, X., & Li, J. (2019). When and how to leverage e-commerce cart targeting: The relative and moderated effects of scarcity and price incentives with a two-stage field experiment and causal forest optimization. *Information Systems Research, 30*(4), 1203–1227.
Ma, S., Tan, H., & Shu, F. (2015). When is the best time to reactivate your inactive customers? *Marketing Letters, 26*, 81–98.
Marcuse, S. (1945). An application of the control chart method to the testing and marketing of foods. *Journal of the American Statistical Association, 40*(230), 214–222.
Meyer, R. J., & Hutchinson, J. W. (2016). (When) are we dynamically optimal? A psychological field guide for marketing modelers. *Journal of Marketing, 80*(5), 20–33.
Montgomery, D. C. (2009). *Introduction to statistical quality control*. John Wiley & Sons.
Neslin, S. A., Taylor, G. A., Grantham, K. D., & McNeil, K. R. (2013). Overcoming the "recency trap" in customer relationship management. *Journal of the Academy of Marketing Science, 41*, 320–337.
Netzer, O., Lattin, J. M., & Srinivasan, V. (2008). A hidden Markov model of customer relationship dynamics. *Marketing Science, 27*(2), 185–204.
Platzer, M., & Reutterer, T. (2016). Ticking away the moments: Timing regularity helps to better predict customer activity. *Marketing Science, 35*(5), 779–799.
Pokornyik, R. (2017). Customer reactivation marketing: Why and how to do it. *Omnisend.* https://www.omnisend.com/blog/customer-reactivation-marketing-why-and-how-to-do-it/.
Reinartz, W., & 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*(4), 17–35.
Reutterer, T., Platzer, M., & Schröder, N. (2021). Leveraging purchase regularity for predicting customer behavior the easy way. *International Journal of Research in Marketing, 38*(1), 194–215.
Robins, J. M., Rotnitzky, A., & Zhao, L. P. (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association, 89*(427), 846–866.
Schmittlein, D. C., Morrison, D. G., & Colombo, R. (1987). Counting your customers: Who they are and what will they do next? *Management Science, 33*, 1–24.
Schwartz, E. M., Bradlow, E. T., & Fader, P. S. (2014). Model selection using database characteristics: Developing a classification tree for longitudinal incidence data. *Marketing Science, 33*(2), 188–205.
Schweidel, D. A., & Fader, P. S. (2009). Dynamic changepoints revisited: An evolving process model of new product sales. *International Journal of Research in Marketing, 26*(2), 119–124.
Seetharaman, P. B., & Chintagunta, P. (2003). The proportional hazard model for purchase timing. *Journal of Business and Economics Statistics, 21*(3), 368–382.
Sharma, S., Niedrich, R. W., & Dobbins, G. (1999). A framework for monitoring customer satisfaction: An empirical illustration. *Industrial Marketing Management, 28*(3), 231–243.
Shewhart, W. A. (1931). *Economic control of quality of manufactured product*. Van Nostrand Company.
Simester, D., Sun, P., & Tsitsiklis, J. N. (2006). Dynamic catalog mailing policies. *Management Science, 52*(5), 683–696.
Stevens, R. (2017). *Ten tips for customer reactivation*. Retrieved September 25, 2018, from https://biznology.com/2017/06/ten-tips-customer-reactivation/.
Van Diepen, M., Donkers, B., & Franses, P. H. (2009). Dynamic and competitive effects of direct mailings: A charitable giving application. *Journal of Marketing Research, 46*, 120–133.

Wager, S., & Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association, 113*(523), 1228–1242.
Wieringa, J. E. (1997). *The case of tin-plating of surface mounted glass diodes*. SOM Research Report 97A18. University of Groningen.
Wieringa, J. E. (1999). *Statistical process control for serially correlated data [Unpublished dissertation]*. SOM Research School: University of Groningen.
Wübben, M., & Von Wangenheim, F. (2008). Instant customer base analysis: Managerial heuristics often "get it right". *Journal of Marketing, 72*(3), 82–93.
Zhang, W., Li, J., & Liu, L. (2021). A unified survey of treatment effect heterogeneity modelling and uplift modelling. *ACM Computing Surveys (CSUR), 54*(8), 1–36.
Zhang, C. W., Xie, M., Liu, J. Y., & Goh, T. N. (2007). A control chart for the Gamma distribution as a model of time between events. *International Journal of Production Research, 45*(23), 5649–5666.
Yoganarasimhan, H., Barzegary, E., & Pani, A. (2022). Design and evaluation of optimal free trials. *Management Science*. In press