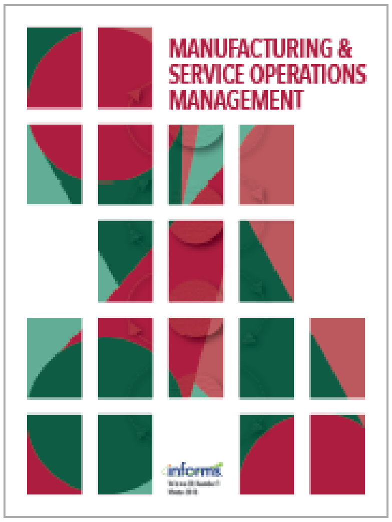


Manufacturing & Service Operations Management

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

Does Customer Email Engagement Improve Profitability? Evidence from a Field Experiment in Subscription Service Retailing

Yiwei Wang, Lauren Xiaoyuan Lu, Pengcheng Shi

MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 24, No. 5, September–October 2022, pp. 2703–2721
ISSN 1523-4614 (print), ISSN 1526-5498 (online)
http://pubsonline.informs.org/journal/msom

Does Customer Email Engagement Improve Profitability? Evidence from a Field Experiment in Subscription Service Retailing

Yiwei Wang,[a] Lauren Xiaoyuan Lu,[b,*] Pengcheng Shi[c]

[a] International Business School, Zhejiang University, Zhejiang, China 314400; [b] Tuck School of Business, Dartmouth College, Hanover, New Hampshire 03755; [c] AI List Capital, Santa Monica, California 90401
*Corresponding author

Contact: yiweiwang@intl.zju.edu.cn, https://orcid.org/0000-0001-6514-1259 (YW); lauren.x.lu@tuck.dartmouth.edu, https://orcid.org/0000-0002-7818-4116 (LXL); kevinshi@ailist.io (PS)

Received: September 15, 2020
Revised: April 16, 2021; July 28, 2021; October 7, 2021
Accepted: December 8, 2021
Published Online in Articles in Advance: July 1, 2022

https://doi.org/10.1287/msom.2022.1122

Abstract. ***Problem definition*:** This paper empirically investigates how customer email engagement affects the profitability of subscription service providers and retailers. They have been using email engagement to increase customer retention. However, it is unclear whether email engagement improves their profitability. The existing literature focuses on email engagement's benefit of customer retention but ignores its associated operating cost to serve retained customers. ***Methodology/results*:** We analyze the outcome of a field experiment conducted by a large U.S. car wash chain that offers tiered subscription services to consumers and employs an radiofrequency identification-based technology to track subscriber service events. We apply survival analysis and difference-in-differences methods to estimate the effects of email engagement on subscribers' retention and service consumption. We find that a one-month engagement with two emails separated by a half-month interval increased the likelihood of subscriber retention by 7.4% five months after the experiment started and decreased the subscriber churn odds by 26.3% for the entire five-month duration. Meanwhile, we find that the same engagement increased a subscriber's per-period service consumption by 7.0%. We provide suggestive evidence for two behavioral mechanisms that explain the effect of email engagement on subscribers' service consumption. First, the engagement effect decays over time and exhibits fatigue after the second email, suggesting that emails act as reminders to subscribers. Second, the engagement effect persists after engagement ends but weakens over time, suggesting the habit formation of subscribers. By computing subscriber lifetime value and the operating cost of service, we find that email engagement increases profit when deployed on mid-level infrequent-use subscribers and top-level subscribers but decreases profit when deployed on mid-level frequent-use subscribers and basic-level subscribers. Therefore, we recommend that the company use a selective strategy by sending engagement emails to only profitable subscribers. ***Managerial implications*:** Our study highlights that email engagement is a double-edged sword; it increases both customer retention and service consumption, and it may decrease profitability when the increased operating cost to serve retained customers outweighs the benefit of customer retention. We recommend that subscription service providers and retailers adopt a data-driven approach to optimize their email engagement strategies.

Supplemental Material: The online appendix is available at https://doi.org/10.1287/msom.2022.1122.

Keywords: consumer behavior • subscription service • retail operations • email engagement • field experiment • data-driven operations

1. Introduction

Subscription, defined as a business model in which customers pay a recurring fee at regular intervals, is an increasingly common way for consumers to buy access to products and services (e.g., health club, curated subscription box service, meal plan, car wash, etc.). This business model has experienced tremendous growth over the past decade, especially in the consumer service and retail sectors. According to a recent report (SUBTA 2019), by 2023, 75% of direct-to-consumer retailers will incorporate a subscription model into their businesses. Moreover, the largest subscription service retailers had more than \$2.6 billion in sales revenue in 2016, which significantly increased from \$57 million in 2011 (Chen et al. 2018). From a firm's perspective, running a subscription model has many benefits, including ensuring a consistent and predictable revenue stream and facilitating personalized interactions as well as improving customer lifetime value and profitability. From a customer's perspective, using a subscription service improves convenience and potentially saves money if she uses the service frequently.

Despite all these benefits, a key challenge for subscription service providers and retailers is customer attrition. Indeed, roughly 40% of subscribers churn within six months of initial enrollment (Chen et al. 2018). To tackle this issue, subscription service providers and retailers employ many strategies to boost customer retention. For instance, they can offer recently churned customers special promotions to motivate them to resubscribe or maintain an interactive online review platform to facilitate more direct communications from and to customers. The predominant and most cost-effective method to improve customer retention is sending emails to customers at regular intervals, which is a key digital engagement strategy in practice (Data & Marketing Association 2015). Engagement emails serve the purposes of both providing information on company activities and reminding customers about their subscribed services. According to a 2015 survey of firms in the consumer retailing and service sectors, 81% of respondents contacted their customers more than twice a month via email in 2014, and 9 of 10 companies declared the strategy of email engagement to be of "great strategic importance" to them (Data & Marketing Association 2015).

As emails and other digital communications have been growing explosively in the last two decades, consumers are now constantly bombarded with marketing emails and text messages, and the effect of email engagement on customer retention has become elusive. According to Data & Marketing Association (2015), 75% of customers resent a brand after receiving excessive engagement emails from the company. Even if email engagement does increase customer retention, its net impact on firm profitability is unclear because engaged customers may increase their service consumption, thereby causing a firm's operating cost to increase. For many consumer services, the marginal cost of providing additional services is substantial compared with the relatively low subscription fee paid by individual customers. The existing literature on email engagement has primarily focused on its benefit of increasing customer retention but ignores its associated operating cost to serve retained customers. Therefore, it is not clear whether customer email engagement will improve the profitability of subscription service providers and retailers.

To fill these gaps in our understanding of email engagement, we seek to answer the following research questions. (1) How does email engagement affect subscribers' retention and service consumption? (2) How should subscription service providers and retailers optimize their email engagement strategies to maximize profitability?

In answering these research questions, we analyze the outcome of a field experiment conducted by our partnering company, a large American car wash chain that offers tiered subscription services to consumers and employs an radiofrequency identification (RFID)-based technology to track subscriber service events. This company owns 130 car wash branches in 16 U.S. states and serves over 168,000 service subscribers nationwide. We apply survival analysis and difference-in-differences (DID) methods to examine the effects of email engagement on subscriber retention and service consumption. The experiment adopts a longitudinal design with email engagement for one month and post-treatment observation for another four months. Our data set is unique for analyzing subscriber behaviors because it contains granular, time-stamped service transaction data collected using RFID devices attached to each subscriber's car. Note that, unlike online platform or e-commerce settings where real-time tracking of individual customer transactions has become widely available, 92.8% of service transactions in the United States still happen in brick-and-mortar facilities (U.S. Department of Commerce 2017), where granular-level data collection is challenging or even infeasible.

Our paper presents several interesting and relevant findings. First, we observe from the field experiment that a one-month engagement with two emails separated by a half-month interval increased the likelihood of subscriber retention by 7.3% five months after the experiment started and decreased subscriber churn odds by 26.3% for the entire five-month duration. Second, we find that the same treatment increased a subscriber's per-period service consumption by 7.0%. Third, we present suggestive evidence for two behavioral mechanisms that explain the effect of email engagement on service consumption. (1) The engagement emails likely acted as reminders to subscribers and increased their service consumption immediately after they received emails, but the engagement effect decayed over time and exhibited fatigue after the second email. (2) The engagement emails increased service consumption even after engagement ends, which is likely because of the habit formation of subscribers.

In sum, these results suggest whether email engagement improves profitability in subscription service settings depends on the relative magnitudes of the engagement effects on subscriber retention and service consumption. This finding stands in sharp contrast to the existing literature on customer engagement, which uses customer retention as the primary outcome measure and mostly finds that email engagement is always beneficial.

Building on our empirical findings, we conduct a data-driven analysis to find the optimal email engagement strategy by computing customer lifetime value and the operating cost of serving subscribers. We find that email engagement increases profit when deployed on all top-level subscribers and mid-level subscribers who infrequently utilized service but decreases profit when deployed on all basic-level subscribers and mid-level subscribers who frequently utilized service. Therefore, we recommend that the company use a selective strategy by sending engagement emails to only profitable subscribers. Our counterfactual analysis estimates that the firm can increase its profit by 17.8% if it adopts this selective engagement strategy. To conclude, our study highlights that email engagement is a double-edged sword for subscription service providers and retailers; it increases both subscriber retention and service consumption, and it may decrease profitability when the increased operating cost to serve retained subscribers outweighs the benefit of subscriber retention. Subscription service providers and retailers need to adopt a data-driven approach to optimize their email engagement strategies.

2. Literature Review

Our work is related to four streams of literature. First, our work naturally falls within the literature of consumer behaviors in response to customer engagement strategies. In various industries, companies have employed various customer engagement strategies, such as emails, customer training, etc., to increase customer retention. Field experiments have been conducted to evaluate the effectiveness of these engagement strategies. For example, Du et al. (2020) implement different engagement strategies to study the effect of text message reminders on the loan repayment rates on a peer-to-peer lending website. Karlan et al. (2016) provide empirical evidence to show that text reminders increase deposits among microfinance customers. Similar to our paper, Charness and Gneezy (2009), Retana et al. (2016), and Calzolari and Nardotto (2017) conduct field experiments to analyze the effect of engagement on customer retention or service consumption. Calzolari and Nardotto (2017) find that sending emails increases service consumptions in a health club. Charness and Gneezy (2009) study the postintervention effects of paying people to attend the gym. They find that providing financial incentives is effective in the formation of healthy habits. Neither of the two papers, however, investigates the effect of emails on customer retention, despite their subscription settings. Retana et al. (2016) document that doing one-shot new customer training can effectively increase short-term customer retention for pay-per-use cloud computing services. However, their analysis does not provide evidence on how customer engagement affects service consumption, a key metric that will affect a subscription service provider's operating cost and profitability. With survival analysis and difference-in-differences analysis, our paper is the first to jointly examine the effects of email engagement on the retention of subscribers and their service consumption behavior. We provide suggestive evidence for two behavioral mechanisms that explain the effect of email engagement on subscribers' service consumption. First, our study empirically demonstrates the patterns of decay and fatigue of the engagement effect that increases service consumption immediately after customers receive emails, which is consistent with the reminder effect documented in the previous literature. Second, we find evidence that email engagement leads to increased service consumption even after engagement ends, which is likely because of the habit formation of subscribers. Although the existing literature primarily focuses on the benefit of customer engagement (e.g., increasing revenue through customer retention), it ignores the costs associated with increased service consumption. Our paper conducts heterogeneous analyses over two important customer characteristics (i.e., the frequency of service consumption and the level of service subscription) in order to evaluate our industry collaborator's email engagement strategy. We recommend that the company use a selective strategy by sending engagement emails to only profitable subscribers.

The second literature our work contributes to is subscription service and retail operations. Operations management researchers have developed various models to study how to manage subscription service and retail operations. For example, Belavina et al. (2017) study the differences in the operational and environmental implications between a subscription model and a pay-per-use model for online grocery delivery. Both Randhawa and Kumar (2008) and Cachon and Feldman (2011) compare the profitability between subscription and pay-per-use models while considering service congestion costs. Danaher (2002) investigates the optimal subscription pricing structure for different cell phone plans. Subscription models have also been examined in information systems under the topic of bundling (Bakos and Brynjolfsson 1999). Unlike informational goods subscription, however, consumer service subscription is usually associated with substantial marginal operating costs. In this paper, we make the first attempt to use data generated from a longitudinal field experiment and use a data-driven approach to assess the trade-off between the benefits of email engagement in improving customer retention and the increased operating costs caused by higher service consumption of engaged customers.

Third, our work contributes to the growing data-driven, practice-based research in operations management. This literature has analyzed a wide range of operational issues in the real world, such as inventory management (Caro and Gallien 2010), pricing (Caro and Gallien 2012, Ferreira et al. 2016, Fisher et al. 2018), information provision (Cui et al. 2019b, Han et al. 2022), and product life cycle (Hu et al. 2019). Within service operations, there have been data-driven works studying delivery service (Cui et al. 2019a, 2021), on-demand service (Bai et al. 2019, Cui et al. 2020), education (Zhang et al. 2017), etc. Our study is an industry-academia collaboration, and our research findings have a direct practical impact on our partnering company. Our data-driven analysis yields an optimal engagement strategy that can potentially increase our partnering company's profit by 17.8%, which demonstrates the real-world relevance of this research.

Fourth, our work is tangentially related to the literature that studies how firms can use innovative technologies to track and study consumer behavior. Many novel technologies, such as RFID, Wi-Fi–based tracking, and mobile targeting, have recently been adopted to study operations management problems in specific industries, such as healthcare (Staats et al. 2017), brick-and-mortar retailing (Hui et al. 2013, Ghose et al. 2019), and e-commerce (Zhang et al. 2019). We complement this literature by showing that large-scale deployment of RFID stickers in a physical setting (specifically, a car wash) is a cost-effective, convenient method to enable a granular analysis of customer service consumption behaviors. Novel data collection technologies, such as RFID stickers, are essential because unlike settings such as an online platform where tracking of customer transactions is common, most service transactions in the United States happen in settings where customer tracking and data collection are still not feasible.

3. Experiment Setting and Hypothesis Development

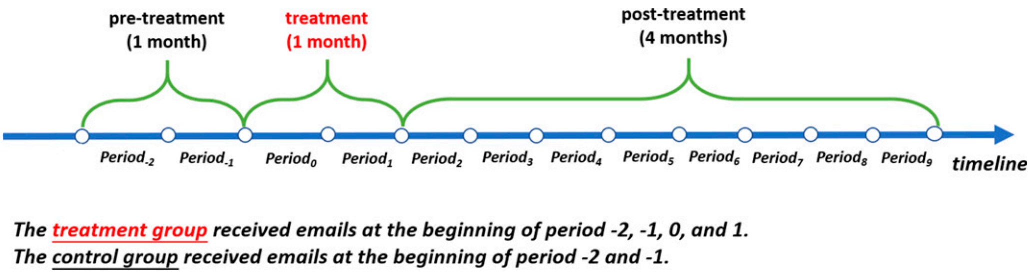
3.1. Experiment Setting

We analyze a field experiment executed at a large U.S. car wash chain. This company operates over 130 drive-through car wash branches located in 16 states and has a customer base of over 168,000 individuals. Figure A3 in the online appendix shows the service coverage of this car wash company. This company mainly operates using a subscription model that offers customers a fixed monthly fee to access uncapped services at any branch operated by this company. An innovative aspect of this company's operations is the use of RFID devices. Specifically, each subscriber is required to attach an RFID sticker underneath her car windshield. The sticker will be immediately destroyed if removed from the vehicle and thus, is not transferable among subscribers. With this novel data collection device, the company's computer system can track each service event's time and location and then, link it to the subscriber's service and renewal history. By partnering with this company, we had obtained a data set that contains all records of consumer activities (e.g., service events, subscription purchases, and subscription status changes) either since January 1, 2016 or since the date the company acquired a local branch, whichever is later.

Our experiment aims to examine the effects of email engagement on customer behaviors. The company intended to achieve two purposes with email engagement: (1) deliver information about recent activities at the company and (2) remind customers about their service subscriptions. In this experiment, all engagement communications were sent through the email channel. Although we cannot provide the actual content of these emails because of the company's restriction, all emails' formats are identical, as illustrated by Figure A1 in the online appendix. In general, an engagement email contains three components: (1) the company's logo, (2) a car wash related picture, and (3) correspondence with subscribers. The specific wordings of such correspondence were generated randomly by the centralized customer relationship management (CRM) software operated by the software as a service company FreshLime, which we also collaborated with for this project. Despite slight variations in wordings, all engagement emails achieve the two purposes mentioned. The CRM software randomizes the assignment of engagement emails to participants. Therefore, the experiment's treatment process was independent of branch characteristics (e.g., our experiment does not suffer from confounding issues that different branches might apply treatments differently or that treatments might be based on subscribers' pretreatment characteristics and transaction history). Furthermore, before the experiment started, customers did not know whether or when they would receive engagement emails.

Figure 1 illustrates the time line of the experiment. Because the participants received each engagement email precisely 15 days after the previous one, we define a 15-day time bucket to be one period.[1] The entire duration of the experiment spanned a total of 12 periods or six months. The first month of the experiment includes period –2 and period –1, which are the pretreatment periods, during which all experiment participants received emails at the beginning of periods –2 and –1. The second month of the experiment includes period 0 and period 1, which are the treatment periods, during which the treatment group received two additional emails at the beginning of periods 0 and 1. The control group no longer received any emails during the treatment periods. From period 2 to period 9, which are the posttreatment periods, neither the treatment group nor the control group

Figure 1. (Color online) Experiment Time Line



received any additional emails, but subscriber renewal and service consumption were tracked during this four-month posttreatment period.

We shall point out that the experiment's starting dates for all participants were staggered, ranging from December 1, 2018 to April 30, 2019, followed by an observation period ending on October 31, 2019. To facilitate our analysis, we create a total of 22 half-month time buckets for the entire 11-month duration of this experiment. The actual experiment starting date for a participant could fall on any date during a half-month time bucket. Figure A2 in the online appendix shows the dates when the customers entered the system and received the treatment.

3.2. Hypothesis Development

Given the experiment setting, we now develop five testable hypotheses to study the effects of email engagements on consumer behavior.

First, as shown in Retana et al. (2016), in the context of information technology, customer engagement activities, such as customer training, can increase satisfaction and loyalty because the customers will better match expectations with specific features of the cloud computing service, resulting in improved customer experiences. The emotional connections developed through the engagement can also increase customer switching costs. As a result, the authors find that customer engagement effectively increases customer retention. In our setting, we examine the impact of customer engagement (i.e., emails) on service subscribers. So, it is plausible that similar effects would appear, and we hypothesize that email engagement would increase subscriber retention.

Hypothesis 1 (Subscriber Retention). *Email engagement increases the retention of a subscriber.*

Second, as shown in Section 2, the literature on customer engagement finds that email engagement can effectively increase customers' service consumption in a wide range of consumer service settings (e.g., banking, health club, vaccine shots, etc.). It is reasonable to hypothesize that email engagement would have a similar impact in our setting, and therefore, the treated subscribers' level of service consumption would increase.

Hypothesis 2 (Subscriber Service Consumption). *Email engagement increases the service consumption of a subscriber.*

Why does email engagement lead to increased service consumption? To answer this question, we resort to past literature's documentation of the reminder effect. According to this theory, an engagement email serves as a behavioral stimulus, which increases the engaged customers' attention to get service (Karlan et al. 2016, Calzolari and Nardotto 2017). Additionally, as reported in a considerable number of psychological studies (Rubin and Wenzel 1996, Baddeley 2007), a memory stimulus (e.g., a reminder) causes temporary peaking of a subject's attention, which would decrease over time unless another stimulus arrives. Therefore, we hypothesize that the engagement effect will decrease over time, as stated in the following hypothesis.

Hypothesis 3 (Engagement Effect: Decay). *A subscriber's service consumption immediately increases upon receiving an engagement email, but this effect decays over time until another email is received.*

In the field experiment, two treatment emails were sent to treated subscribers at a 15-day interval. This design allows us to test how subscribers respond to sequential email engagements. Researchers find when individuals become accustomed to recurring stimuli, they become desensitized and less responsive to future stimuli they receive (Boksem and Tops 2008, Calzolari and Nardotto 2017). As a result, the first stimulus's effect is likely more significant than subsequent ones. So, we conjecture that the second treatment email's effect on subscribers' consumption would be less significant than the first treatment email in our setting, as stated in the following hypothesis.

Hypothesis 4 (Engagement Effect: Fatigue). *The positive effect of the second treatment email on a subscriber's service consumption is smaller than that of the first treatment email.*

In our setting, email engagement lasted for two months (including both pretreatment and treatment periods). According to the literature, after customers maintain a high level of service consumption for a considerably long period, habituation may happen (Charness and Gneezy 2009, Calzolari and Nardotto 2017). As a result, the increased service consumption would persist even after email engagement ends. However, the positive postengagement effect on service consumption may weaken over time without continued email engagement, as evidenced by a previous study on users' gym attendance (Calzolari and Nardotto 2017).

Hypothesis 5 (Postengagement Effect). *The positive effect of email engagement on a subscriber's service consumption persists even after engagement ends, but this effect weakens over time.*

4. Data Description

4.1. Sample Construction

To systematically examine the effects of email engagement on subscriber retention and service consumption, our industry collaborator conducted a field experiment that started on December 1, 2018 and finished on October 31, 2019. During this experimental period, 4,393 new customers were selected to participate in an email engagement program initiated by the company. To be eligible to participate, customers must first enroll in the company's subscription program online and get an RFID sticker to place on their windshield at any of its local branches. Then, a fixed recurring monthly fee will be deducted from a subscriber's credit card on file until she cancels the subscription. The participants' average subscription tenure at the start of the experiment was less than half a month, whereas the maximum subscription tenure was two months, which means these were newly enrolled customers.

Furthermore, each customer must enroll in one of the three subscription programs: basic-level, mid-level, and top-level. Basic-level subscribers pay a monthly fee between \$10 and \$20, mid-level subscribers pay a monthly fee between \$20 and \$30, and top-level subscribers pay a monthly fee between \$30 and \$40. Table A1 in the online appendix shows the service options included in different subscription levels. Note that the service options are identical throughout the entire car wash chain, whereas the monthly subscription fee may differ slightly across branches. Among the participants, we exclude the following subscribers from our analysis: (1) those who canceled subscriptions before the treatment started and (2) promotion subscribers who paid no subscription fees. This exclusion reduces our final sample to 4,077 customers, among whom 1,435 (35.2%) customers chose the basic-level program, 1,763 (43.2%) chose the mid-level program, and 879 (21.6%) chose the top-level program. Table 1 provides the definition and summary statistics of key variables used in our study.

Among all experiment participants, 1,626 (40%) were assigned to the control group, and 2,451 (60%) were assigned to the treatment group. As a balance check, we confirm that both treatment and control groups were comparable along with a set of important pretreatment variables, including the number of distinct branches visited before treatment, the subscription tenure before treatment, service consumption in periods –2 and –1, and state average temperature. A nonparametric test indicates that all *p*-values for comparing these variables between the treatment and control groups are above 0.171, thus statistically insignificant. Table A2 in the online appendix shows the result of this balance check. Since the treatment began, only a small fraction of customers (<4%) upgraded their subscription levels, and we find no evidence to prove that service upgrade behavior correlates with the treatment applied in our experiment.

Table 1. Summary Statistics

Variable	Definition	Observations	Mean	Standard deviation
Basic-level subscribers (35.2%)				
Treat	1 if a subscriber received engagements and 0 otherwise	1,435	0.53	0.50
Distinct Branches Visited	Number of distinct stores visited prior to treatment	1,435	1.34	0.63
Subscription Tenure	Total months of subscription prior to treatment	1,435	1.52	0.40
Total Prior Consumption	Service consumption (visits) in period_{-2} and period_{-1}	1,435	2.87	2.33
Monthly Subscription Fee	Monthly fee paid by each subscriber	1,435	14.90	1.53
Mid-level subscribers (43.2%)				
Treat	1 if a subscriber received engagements and 0 otherwise	1,763	0.64	0.48
Distinct Branches Visited	Number of distinct stores visited prior to treatment	1,763	1.37	0.68
Subscription Tenure	Total months of subscription prior to treatment	1,763	1.46	0.38
Total Prior Consumption	Service consumption (visits) in period_{-2} and period_{-1}	1,763	3.15	2.73
Monthly Subscription Fee	Monthly fee paid by each subscriber	1,763	24.59	3.08
Top-level subscribers (31.6%)				
Treat	1 if a subscriber received engagements and 0 otherwise	879	0.64	0.48
Distinct Branches Visited	Number of distinct stores visited prior to treatment	879	1.35	0.66
Subscription Tenure	Total months of subscription prior to treatment	879	1.46	0.39
Total Prior Consumption	Service consumption (visits) in period_{-2} and period_{-1}	879	3.54	3.40
Monthly Subscription Fee	Monthly fee paid by each subscriber	879	35.66	2.39

4.2. Dependent Variables

We observe each customer's service consumption and subscription renewals for the entire six-month duration of the experiment, as illustrated in Figure 1. Two main dependent variables we consider are subscriber retention and service consumption.

The subscriber retention variable captures whether a participant was renewing her service subscription by the end of each period. Given our data set, identification of a customer churn event is straightforward. Namely, if a subscriber has not renewed the service in any given month, she is marked as churned by the end of that month. The specific date of churn is then defined to be the subscriber's membership expiration date, which is one month after the date of her last subscription renewal event. Only a small fraction of subscribers (<2%) marked as churned would resume subscription at a later date during the experiment, and we find no evidence that such behavior is systematically associated with the assignment of treatment. For this group of subscribers, we identify the churn period to be the period after the first churn event occurs. Using alternative identification approaches (e.g., the period after the last churn event) or removing these customers does not qualitatively change our results. In our sample, 30.5% of participants had churned by the end of the experiment (i.e., period 9).

The service consumption variable represents the total number of visits each subscriber made in any period. It is worth noting that email engagement may affect a subscriber's service consumption through the direct effect of increasing her per-period consumption or indirectly by reducing her churn rate. Our data set's unique feature is that if a customer has no consumption in a period, we can clearly identify whether it is because of her subscription cancellation or service inactivity. Thus, in this paper, we focus on the direct effect of email engagement on a subscriber's service consumption.

4.3. Key Control Variables

For our survival analysis, we include two sets of control variables. The first set of variables is related to pretreatment subscriber characteristics, including subscription tenure, subscription expense, distinct branches visited prior to the treatment date, consumption in period –1, consumption in period –2, and the subscription level (i.e., basic-level, mid-level, or top-level). We denote this set of variables by X_i.

The second set of control variables is related to the timing of treatment for each participant. Specifically, we employ 12 dummy variables $StartingPeriod_{ni}$ to represent the starting time of the experiment ranging from December 1, 2018 to April 30, 2019 whose value is equal to one if the email engagement for subscriber i started in period n. We use two binary dummy variables, $Email1Weekday_i$ and $Email2Weekday_i$, to control for whether subscriber i received the two engagement emails on a weekday or a weekend during the treatment duration. In our sample, all experiment participants survived through the pretreatment periods (i.e., periods –2 and –1), so it is unnecessary to add control variables for the pretrend of customer retention behavior. Similar time-related control variables have been previously used for survival analysis in the literature (see, e.g., Retana et al. 2016).

For our service consumption analysis, we will leverage our data set's panel structure by including subscriber fixed effects to control for potential pretreatment heterogeneity at the subscriber level. With subscriber fixed effects, it is unnecessary to include other time-invariant control variables at the subscriber level. To control for heterogeneous branch characteristics, we include branch fixed effects. It is feasible to include branch fixed effects in our setting because the relationship between branches and subscribers is not hierarchical, and a significant portion of subscribers in our sample visited multiple branches. To capture time-variant, branch-level heterogeneity, we also include a set of time-variant control variables denoted by Z_{jt-1}, which represents the lagged aggregate service consumption by subscribers and pay-per-use customers at branch j in period $t-1$.

5. Empirical Methods

Our empirical strategy employs two main methods. First, we use survival analysis to investigate the effect of email engagement on subscriber retention. Second, we adopt the difference-in-differences method to analyze the effect of email engagement on service consumption.

5.1. Subscriber Retention

We employ linear probability and probit models to estimate the effect of email engagement on subscriber retention. The dependent variable is $Survival_i$, indicating whether subscriber i "survived" by the end of the experiment (i.e., period 9). The specifications for the linear probability and the probit models are given as follows:

$$Survival_i = \alpha_0 + \alpha_1 Treat_i + \alpha_2 X_i + \alpha_3 Time_i + \epsilon_i, \quad (1)$$

$$Pr(Survival_i) = \Phi(\alpha_0 + \alpha_1 Treat_i + \alpha_2 X_i + \alpha_3 Time_i + \epsilon_i). \quad (2)$$

$Treat_i$ is a binary variable, which equals one if subscriber i received an additional month of email engagement and zero otherwise. The coefficient α_1 captures the treatment effect, which will be positive if email engagement increases subscriber retention. X_i is the set of control variables that capture pretreatment subscriber characteristics, and $Time_i$ is the set of control variables related to the timing of the experiment, both of which are described in Section 4.3.

Other than linear probability and probit models, we employ the logit hazard model to analyze email engagement's effect on subscriber churn. The logit hazard model has been widely adopted in longitudinal data analysis (Singer et al. 2003). To apply this model, we use a binary outcome variable $Churn_{it}$ to indicate how likely subscriber i will churn in period t. For subscriber i, these variables equal zero for all periods before the churn event occurs, equal one for the period when the churn event occurs, and equal null for periods after the churn event. Our model is specified as follows:

$$\log\left(\frac{p(Churn_{it}=1)}{1-p(Churn_{it}=1)}\right)=\sum_{t=0}^{9}\beta_t D_t+\alpha_1 Treat_i+\alpha_2 X_i +\alpha_3 Time_i+\epsilon_i. \tag{3}$$

This logit hazard model uses a logit function to "link" all explanatory variables on the right-hand side of this equation to the outcome variable $Churn_{it}$. Thus, the term on the left-hand side of this equation represents the log hazard odds of the churn event. D_t is the indicator variable that equals one for period t and zero otherwise. The coefficient β_t represents the underlying baseline hazard that all subscribers are subject to in period t. The treatment effect is captured by the coefficient α_1, which will take a negative value if the treatment has a positive effect on churn reduction. With this model, $e^{\alpha_1}-1$ corresponds to the percentage change of the hazard odds (i.e., the probability of churn over the probability of being retained) for the treatment group. X_i and $Time_i$ are the sets of control variables identical to those used in our linear probability and probit models, and ϵ_i is the error term. It should also be noted that the logit hazard model is similar to the regular logit model, and therefore, we can use the standard maximum likelihood method to estimate it.

5.2. Service Consumption

To estimate the effect of email engagement on service consumption, we apply the DID model with count data and conduct a Poisson regression. The DID model has been widely used in economics for policy evaluations (see, e.g., Duflo 2001) and recently adopted in the empirical service operations literature (Cui et al. 2019a, 2021). In our model, the dependent variable C_{ijt} is the total number of service visits of subscriber i made at branch j in period t. The count data model (i.e., Poisson) is appropriate for our setting as the number of visits only takes nonnegative integer values. Recall from Section 4.2 that we focus on the effect of email engagement on service consumption conditional on a subscriber being retained. To *exclude* the effect of customer churn on service consumption, we remove observations of churned customers during their postchurn periods. To account for the possible issue of overdispersion of zero entries, we use the robust variance-covariance matrix for our Poisson maximum likelihood estimator (Retana et al. 2016). Our baseline DID model is specified as

$$\log(E[C_{ijt}])=\beta_0+\beta_1 Treat_i\times Post_{it}+Post_{it}+\beta_2 Z_{jt-1} +\mu_i+\gamma_j+\theta_t+\epsilon_{ijt}. \tag{4}$$

Our observation unit is a service transaction of a subscriber in a branch. Here, i denotes each subscriber, j denotes a specific car wash branch, t denotes the period number, μ_i denotes subscriber fixed effects, γ_j represents branch fixed effects, θ_t captures the period fixed effect, and ϵ_{ijt} is the error term. This three-way fixed effects model[2] is feasible in our setting because the relationship between branches and subscribers is not hierarchical and because subscribers may visit multiple branches. $Treat_i$ is a binary variable that equals one for the treatment group and zero for the control group. $Post_{it}$ equals zero for the first two periods (i.e., periods −2 and −1) and one for all periods (including two treatment periods and eight posttreatment periods; i.e., period 0 through period 9) after the treatment started. We incorporate the term $Post_{it}$ in our specification to account for subscriber-specific time trends. For the control group, $Post_{it}$ is also well defined even for subscribers in the control group because we know which periods are their pretreatment periods and thus, find their corresponding "posttreatment" periods. The coefficient β_1 captures the treatment effect of the two engagement emails sent during the treatment period on subscribers' service consumption.

Note that the dependent variable is a count variable, and the Poisson regression model specifies the log of the expected count as a function of the predictive variables (Wooldridge 2010). So, the coefficient β_1 can be interpreted as follows; with email engagement, the log of expected service consumption increases by β_1. In other words, given email engagement, the percentage change in the expected service consumption is $e^{\beta_1}-1$. In our model, we observe all participants for exactly 12 periods (i.e., 2 pretreatment periods, 2 treatment periods, and 8 posttreatment periods). The matrix Z_{jt-1} contains two vectors, which capture the total service consumption for all subscribers or pay-per-use customers at each branch j in period $t-1$. Note that we can only obtain the consumption information for pay-per-use customers from the point-of-sale (POS) data because these customers were not equipped with the RFID tracker. The POS data were aggregated at the branch-period level instead of the subscriber-period level. These time-variant branch-level control variables were lagged by one period to avoid the issue of reverse causality. Finally, standard errors are clustered at the subscriber level instead of at the branch level because subscribers may visit multiple branches.

To explore the decay and fatigue patterns of the email engagement effect on service consumption (i.e., Hypotheses 3 and 4), we conduct a second DID analysis focusing on a short 37-day time frame (i.e., 7 days before and 30 days after the treatment). For this analysis, the panel data are constructed at the daily level, and we label the date any subscriber received her first treatment email as day 0 (for the control group, we also label the matching treatment date, despite the fact that no email was dispatched). Because of the daily-level panel structure, for this regression, our dependent variable is a binary variable $Service_{it}$, indicating whether a subscriber i had service or not on day t. We do not use the number of visits as our dependent variable, as it is unlikely that customers will get two car washes within a day. In addition, we conduct this regression at the subscriber level instead of the subscriber-branch level because it is unlikely that a subscriber would seek service at multiple branches in a single day. With a binary dependent variable, we adopt the following logistic regression for our DID analysis to study the effect of email engagement on the probability of service consumption at the subscriber-day level:

$$\begin{aligned}Pr(Service_{it}) = {} & \beta_0 + \beta_1 Treat_i \times Day_{-1it} + \beta_2 Treat_i \\ & \times Day_{0it} + \beta_3 Treat_i \times Day_{1\sim7it} + \beta_4 Treat_i \\ & \times Day_{8\sim14it} + \beta_5 Treat_i \times Day_{15it} + \beta_6 Treat_i \\ & \times Day_{16\sim22it} + \beta_7 Treat_i \times Day_{23\sim29it} + \alpha_1 X_i \\ & + \alpha_2 Time_i + Day_{-1it} + Day_{0it} + Day_{1\sim7it} \\ & + Day_{8\sim14it} + Day_{15it} + Day_{16\sim22it} \\ & + Day_{23\sim29it} + Day_t + \epsilon_{it}. \end{aligned} \tag{5}$$

$Treat_i$ is a binary variable, which equals one if subscriber i is in the treatment group. Day_{-1it}, Day_{0it}, etc. are dummy variables equal to one for the corresponding time bucket and zero otherwise. Coefficients β_2 and β_5 capture the effect of email engagement on days 0 and 15 when the first and second treatment emails were dispatched. Coefficients β_3 and β_4 capture the effect of email engagement on the average daily probability to get service during the two weeks following the dispatch of the first treatment emails (i.e., day 1 to day 7 and day 8 to day 14); β_6 and β_7 capture the effect of email engagement on the average daily probability to get service during the two weeks following the dispatch of the second treatment email (i.e., day 16 to day 22 and day 23 to day 29). X_i is the set of control variables that capture pretreatment subscriber characteristics, and $Time_i$ is the set of control variables related to the experiment's timing, both of which are described in Section 4. Day_t is the day fixed effect (except for day –1, day 0, and day 15). The coefficient ϵ_{it} is the error term.

Finally, to investigate the postengagement effect (i.e., Hypothesis 5), we conduct additional regression analyses to explore the dynamic, long-term effects of email engagement on service consumption. Recall that each period in our experiment contains 15 days. To have an intuitive interpretation for the postengagement effect, we study the dynamics at the month level (i.e., two consecutive periods are referred to as a month, although these months do not have to coincide with the calendar months). We run the regression analysis given by the following model:

$$\begin{aligned}\log(E[C_{ijt}]) = {} & \beta_0 + \beta_1\ Treat_i \times DuringTreatMonth0_{it} \\ & + \beta_2\ Treat_i \times PostTreatMonth1{,}2_{it} \\ & + \beta_3\ Treat_i \times PostTreatMonth3{,}4_{it} \\ & + \beta_4 Z_{jt-1} + DuringTreatMonth0_{it} \\ & + PostTreatMonth1{,}2_{it} \\ & + PostTreatMonth3{,}4_{it} \\ & + \mu_i + \gamma_j + \theta_t + \epsilon_{ijt}, \end{aligned} \tag{6}$$

where C_{ijt} represents the number of services visits subscriber i received at branch j in period t. $DuringTreatMonth0_{it}$, $PostTreatMonth1{,}2_{it}$, and $PostTreat\ Month3{,}4_{it}$ are dummy variables that capture dynamics of the treatment. Specifically, $DuringTreatMonth0_{it}$ equals one for month 0 (i.e., periods 0–1) when the treatment was being applied and equals zero otherwise. $PostTreat\ Month1{,}2_{it}$ equals one for posttreatment months 1 and 2 (i.e., periods 2–5) and equals zero otherwise. $PostTreat\ Month3{,}4_{it}$ equals one for posttreatment months 3 and 4 (i.e., periods 6–9) and equals zero otherwise. To this end, the coefficients β_1, β_2, and β_3 capture the treatment effect for associated time buckets. Similar DID specifications have previously been adopted to study the long-term effect of sudden removal and restoration of high-quality delivery options for an e-commerce retail platform (Cui et al. 2019a).

5.3. Identification

This section discusses potential issues related to identifying the causal relationship between email engagement and consumer behaviors. Causal inference has been a notoriously difficult empirical question because of endogeneity problems, such as self-selection and unobserved heterogeneity (Rubin 1974). Field experiments, however, provide a clean way to identify causal effects, overcoming potential confounding factors that may result in biased estimations of the actual treatment effect. In this research, we exploit a controlled experiment setting, where the treatment of interest is the two emails an experiment subject received in periods 0 and 1. In the following, we will show that this exogenous intervention is sufficient to allow us to identify the causal effect of email engagement in our context.

First, in our experiment, the treatment application process was managed by a centralized CRM software system. In particular, new subscribers from all car wash branches were placed in a first come, first served queue and then randomly assigned to either the treatment group or the control group. According to Rubin (1974), proper randomization helps establish the comparability of treatment and control groups. We validate the randomization process's effectiveness by conducting a balance check between the treatment group and the control group. The results are shown in Table A2 in the online appendix. According to Table A2 in the online appendix, there is no statistically significant difference between the treatment group and the control group for pretreatment subscriber characteristics, such as distinct branches visited, subscription tenure, pretreatment consumptions, and state average temperature.

Moreover, before the treatment commenced, all experiment participants were unaware of the total number of emails they would receive, nor did they know the frequency of those emails. Therefore, our setting is free from the self-selection bias (i.e., when subjects select themselves into a group, resulting in a biased sample).

For this experiment, we track when each subscriber received an engagement email. However, we do not observe whether she opened the email. So, our analysis is focused on the notion of *intention to treat* (e.g., Rubin 1974) by studying all customers who received engagement emails. The efficacy of intention to treat is of primary interest to our industry collaborator because a subscription service provider can only control intention to treat but not directly control whether engagement emails are viewed. Hence, the issue of whether subscribers viewed those emails is beyond the scope of this paper.

One potential concern in our experiment is unobserved intertemporal and cross-sectional heterogeneity, which may arise because each participant received engagement emails in different time periods and consumed services at different branches. These differences can potentially correlate with the outcomes and yield a biased estimate of the average treatment effect. For the service consumption analysis, we address this issue by using the DID approach, which identifies the causal effect by relying on the within-subscriber variation across time. In the DID specification, we control for pretreatment heterogeneity of subscriber characteristics by including subscriber fixed effects and control for branch-level heterogeneity by including branch fixed effects. In addition, we include period fixed effects to account for intertemporal heterogeneity. Finally, we use branch-level aggregate service consumption to control for each branch's time-variant system congestion level. To avoid the reverse causality between the dependent variable of subscriber service consumption and the control variable of per-period aggregate service consumption, we use one-month lagged variables for aggregate service consumption. Controlling for these time-variant covariates can effectively mitigate the problem of unobserved heterogeneity. For the survival analysis and the daily-level consumption analysis, we include a set of variables to control for heterogeneous subscriber characteristics and each subscriber's experiment starting time, as described in Section 3.

It should be noted that our industry collaborator designed and conducted this field experiment without stratified sampling. So, the fractions of customers enrolled in different subscription levels are not equal between the treatment and control groups. As an additional robustness check, we use matching algorithms to create 1,620 treatment-control pairs from the full sample and reestimate the treatment effect using the matched sample. To do so, we first perform an exact match between a pair of control and treatment customers according to their service level and the treatment starting period. Then, we employ the nearest distance matching algorithm with Mahalanobis distance to create "equivalent" treatment-control pairs. That is, for each treatment-control pair ij, the following expression is minimized:

$$D_{ij} = \sqrt{(X_i - X_j)' S^{-1} (X_i - X_j)}. \quad (7)$$

Vector X contains a customer's pretreatment characteristics, including subscription tenure at the time of treatment, the number of unique branches visited, and pretreatment service consumptions. S^{-1} is the covariance matrix between each of customer i and j's pretreatment characteristics. After matching, for both the treatment and control groups, the percentage of customers in each service level is identical, and the total number of customers who started the experiment in each period is identical. We use this matched sample to conduct a robustness check. Estimation results obtained using the matched sample are quantitatively similar to those obtained using the full sample (see Tables A3 and A4 in the online appendix).

6. Empirical Results

We present the estimation results for the effects of email engagement on subscriber retention in Section 6.1 and the effects on service consumption in Section 6.2. Section 6.3 explores the mechanisms through which email engagement influences consumer behavior.

6.1. The Effect of Email Engagement on Subscriber Retention

We first visually inspect the effect of email engagement on subscriber retention. Figure 2 shows the cumulative survival probability for the treatment and control groups by the end of each period. At the beginning of periods –2 and –1, the first and second engagement emails were sent to all participants in the experiment, and the corresponding survival probabilities for both treatment and control groups were 100%, as our sample is constructed to include only participants who survived at least through the treatment starting date. At the beginning of periods 0 and 1, two additional engagement emails were sent to the treatment group but not the control group. Figure 2 shows that starting from period 0, the survival probabilities for both groups started to decline, but that of the treatment group constantly stayed above that of the control group, suggesting a positive subscriber retention effect because of the treatment of the two additional engagement emails.

To quantify email engagement's effect on subscriber retention, we estimate both a linear probability and a probit model. Columns (1)–(3) in Table 2 report the estimation results. To help interpret the magnitude of the coefficients, we report the corresponding marginal effect in the bottom section of the table. Results from both regressions are qualitatively and quantitively similar and statistically significant. Column (1) shows that under the linear probability model, email engagement increased the likelihood of subscriber survival through period 9 by 7.4%. The probit model yields a similar estimation, as shown in column (2). The estimated effect of email engagement on the subscriber churn hazard rate is presented in column (3). The coefficient of $Treat_i$ is −0.305 and statistically significant, indicating that the treatment group had a lower churn rate. Moreover, this coefficient translates to a 26.3% reduction in the hazard odd, which indicates that the ratio of the probability of a subscriber canceling her service to the probability of the subscriber retaining her service is reduced by 26.3% by period 9, which is five months after the treatment started. Altogether, these results imply that email engagement effectively increased subscriber retention (or decreased customer churn). These results support Hypothesis 1.

In addition to the main regression analysis, we explore the heterogeneity in the effects of email engagement using two important customer characteristics: subscription level and pretreatment consumption frequency. These are two key dimensions that capture customer preference and behavioral patterns in our setting. A customer's subscription level reflects her self-selected preference of service level, whereas a customer's consumption frequency reveals her actual pattern of service consumption.

The heterogeneous analysis results for pretreatment consumption frequency are reported in Table 2, columns (4)–(9). In the field experiment, all participants were new subscribers, and the maximum subscription tenure before treatment was two months. Hence, we use the median of the total pretreatment consumption to classify customers into two groups: infrequent users (i.e., less than or equal to four visits) and frequent users (more than four visits). We then conduct linear probability, probit, and logit hazard models for each subsample. According to columns (4)–(9), email engagement reduced the hazard odds by 20.5% for infrequent users and 35.3% for frequent users. Both estimates are statistically significant. Therefore, email engagement has a stronger retention effect on frequent users than on infrequent users.

The heterogeneous analysis results for subscription levels are reported in Table 3. According to columns

Figure 2. (Color online) Subscriber Survival Curve: All Subscribers

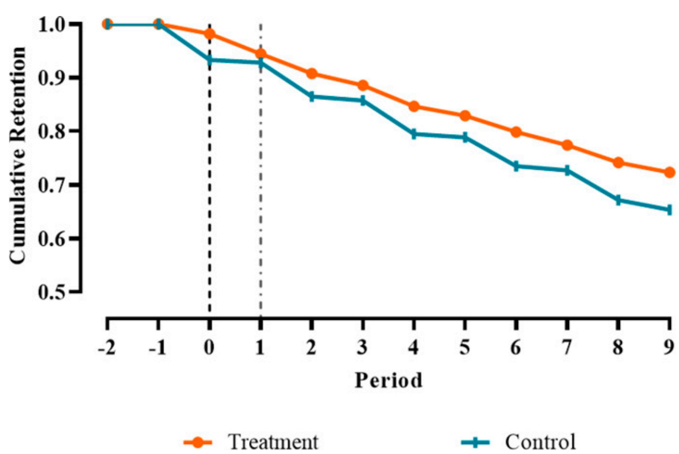


Table 2. The Effect of Email Engagement on Subscriber Retention: Full Sample, Frequent and Infrequent Users

	Samples								
	Full sample			Infrequent users			Frequent users		
Dependent variables	*Survival LPM* (1)	*Survival Probit* (2)	*Churn Logit Hazard* (3)	*Survival LPM* (4)	*Survival Probit* (5)	*Churn Logit Hazard* (6)	*Survival LPM* (7)	*Survival Probit* (8)	*Churn Logit Hazard* (9)
Treat	0.074***	0.213***	−0.305***	0.061**	0.167**	−0.230**	0.091***	0.278***	−0.436***
	(0.016)	(0.047)	(0.066)	(0.023)	(0.064)	(0.087)	(0.023)	(0.072)	(0.103)
Marginal Effect	0.074***	0.073***		0.061**	0.059**		0.091***	0.091***	
Δ Hazard Odds, %			−26.3***			−20.5**			−35.3***
Subscriber characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,077	4,077	4,077	2,186	2,186	2,186	1,891	1,891	1,891

Note. Standard errors are given in parentheses.
$^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

(3), (6), and (9), email engagement reduced the hazard odds by 25.2%, 16.8%, and 40.7% for basic-level, mid-level, and top-level subscribers, respectively. All estimates are statistically significant. In sum, email engagement has the strongest retention effect on top-level subscribers and the weakest retention effect on mid-level subscribers. Interestingly, the effect of email engagement on retention is not monotone with respect to subscription levels. Top-level subscribers self-selected themselves to the highest subscription level likely because they are the most loyal customers or because they value the subscription the most. It turns out these loyal customers were retained the most through email engagement. Why basic-level subscribers were retained more than mid-level subscribers can be explained by how email engagement changed their service consumption behavior. As we will show in Section 6.2 next, basic-level subscribers had the most percentage increase in their service consumption among all subscribers. Given their behavioral changes, basic-level subscribers might perceive their subscriptions more valuable than prior to email engagement and thus, were more retained than mid-level subscribers.

6.2. The Effect of Email Engagement on Service Consumption

In this section, we examine the effect of email engagement on subscribers' service consumption. Figure 3 shows the average service consumption for the treatment and control groups during each period. In the pretreatment periods (i.e., periods –2 and –1), the pretrends of the treatment and control groups are almost identical to each other, which supports the parallel pretrend assumption of our DID specification. During and after the treatment periods (i.e., periods 0 and 1), the control group's service consumption declined much more quickly than the treatment group. Moreover, this effect was persistent and lasted through the end of period 9.

After inspecting Figure 3, we turn to the DID regression results of service consumption using the full

Table 3. The Effect of Email Engagement on Subscriber Retention: Different Subscription Levels

	Samples								
	Basic-level subscribers			Mid-level subscribers			Top-level subscribers		
Dependent variables	*Survival LPM* (1)	*Survival Probit* (2)	*Churn Logit Hazard* (3)	*Survival LPM* (4)	*Survival Probit* (5)	*Churn Logit Hazard* (6)	*Survival LPM* (7)	*Survival Probit* (8)	*Churn Logit Hazard* (9)
Treat	0.064**	0.199*	−0.290*	0.046†	0.135†	−0.184†	0.145***	0.387***	−0.522***
	(0.026)	(0.078)	(0.116)	(0.025)	(0.071)	(0.100)	(0.036)	(0.096)	(0.122)
Marginal Effect	0.064**	0.064*		0.046†	0.049†		0.145***	0.146***	
Δ Hazard Odds, %			−25.2*			−16.8†			−40.7***
Subscriber characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,435	1,435	1,435	1,763	1,763	1,763	879	879	879

Note. Standard errors are given in parentheses.
$^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Table 4. The Effect of Email Engagement on Service Consumption

	Samples					
	Full (1)	Infrequent users (2)	Frequent users (3)	Basic-level subscribers (4)	Mid-level subscribers (5)	Top-level subscribers (6)
Treat × *Post*	0.068**	0.122**	0.064*	0.137***	0.108**	0.065
	(0.026)	(0.046)	(0.031)	(0.038)	(0.037)	(0.053)
Branch characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Branch fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Subscriber fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	78,231	35,497	42,734	28,442	33,903	15,886
Log likelihood	−81,080	−32,332	−48,499	−29,228	−34,622	−17,594

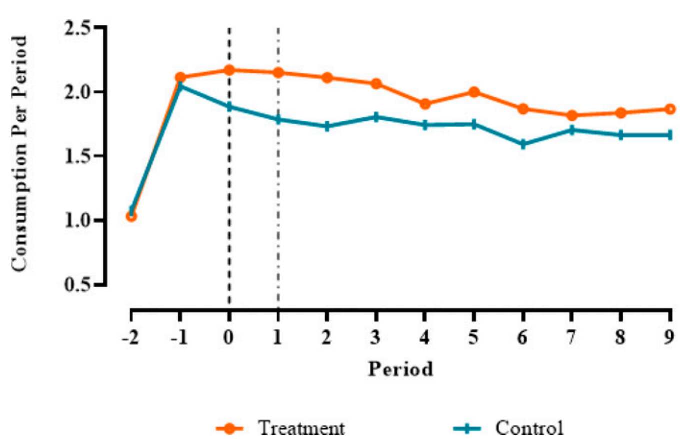
Note. Standard errors are given in parentheses.
$^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

sample, reported in Table 4, column (1). Note that there were 4,077 subscribers in the full sample. However, 55 were dropped from the DID analysis because they did not obtain any service in any period despite paying the subscription fee. This reduces the total number of subscribers to 4,022. Each subscriber might visit more than one car wash branch, so our regression is conducted at the subscriber-branch level. Also, we construct our panel data to be unbalanced as each subscriber-branch pair has an observation for a period only if the subscriber renews subscription through that period. Column (1) shows that the treatment group's consumption increase is positive and statistically significant with a magnitude of 7.0% ($=e^{0.068} - 1$). This result supports Hypothesis 2.

We next report the heterogeneous treatment effects for frequent and infrequent users. Columns (2) and (3) of Table 4 represent the treatment effect on service consumption for infrequent and frequent users. For infrequent users, email engagement increased their consumption by 13.0% or 0.31 visits in absolute terms. For frequent users, email engagement increased their consumption by 6.6% or 0.55 visits in absolute terms. The results of both regressions are statistically significant.

Finally, we conduct the heterogeneous analysis of service consumption for different subscription service levels. For basic-level, mid-level, and top-level subscribers, email engagement increased their consumption by 14.7%, 11.4%, and 6.7%, respectively. The estimation results are statistically significant for basic-level and mid-level subscribers. Taken together, the increase of service consumption is the smallest for top-level subscribers, medium for mid-level subscribers, and the largest for basic-level subscribers. To understand the different engagement effects on the service consumption of different subscribers, note that top-level subscribers on average consumed service 3.54 times per month prior to treatment. This is close to having a car wash every week. Hence, increasing the car wash frequency further would be difficult. The email engagement's effects on top-level

Figure 3. (Color online) Subscriber Service Consumption Curve: All Subscribers



subscribers are therefore the smallest. Following a similar logic, the basic-level subscribers on average consumed service 2.87 times per month prior to treatment, which is the lowest among all subscribers. Therefore, there is a lot of room for them to increase their service frequency.

6.3. Evidence on Mechanisms

This section explores behavioral mechanisms through which email engagement led to the observed increase in service consumption. Specifically, we present evidence on the patten of decay and fatigue of the engagement effect on service consumption in Section 6.3.1. We investigate the postengagement effect on service consumption in Section 6.3.2.

6.3.1. The Engagement Effect: Decay and Fatigue. To explore the decay and fatigue patterns of email engagement's effect on service consumption, we conduct a DID analysis (with a 37-day time frame) on service consumption according to Equation (5). The estimations results are presented in Table 5. We first discuss the decay of the engagement effect. As expected, there is no statistically significant difference in the probability of obtaining service ($\beta_1 = -0.048$) between the treatment and control groups the day before treatment (day –1). On day 0, when the first treatment email was sent, the treatment effect is significant and positive ($\beta_2 = 0.373$), which translates to a 37.3% increase in the daily consumption probability. From day 1 to day 7, the treatment effect is still positive and significant ($\beta_3 = 0.144$) but decreased to a 14.4% increase in the daily consumption probability. The estimated DID coefficient from day 8 to day 14 ($\beta_4 = 0.149$) is quantitatively close to the estimation for the week before. This decay of the positive engagement effect within two weeks of receiving engagement emails supports Hypothesis 3.

We next discuss the fatigue of the engagement effect because of repeated email engagements. On day 15, the second treatment email was sent to treated subscribers. For that day, the treatment effect was positive but insignificant ($\beta_5 = 0.116$). For day 16 to day 22, a significant increase in the daily probability to get service emerged again ($\beta_6 = 0.093$), although at a much smaller magnitude than that for day 1 to day 7. For day 23 to day 29, the estimated DID coefficient was significant at $\beta_7 = 0.106$. If we compare the estimations of DID coefficients of the first and second treatment emails, we observe that the increase in consumption probability is stronger for the first treatment email than that for the second treatment email in any time bucket (including the treatment day, the first week after treatment, and the second week after treatment) throughout the 37-day time frame. These observations support Hypothesis 4 that the engagement effect exhibits a pattern of fatigue for the second email.

Table 5. The Decay and Fatigue of the Effect of Email Engagement on Service Consumption

	Average Daily Probability to Get Service
Treat × Day −1	−0.048
	(0.107)
Treat × Day 0	0.373***
	(0.114)
Treat × Day 1–7	0.144**
	(0.055)
Treat × Day 8–14	0.149**
	(0.055)
Treat × Day 15	0.116
	(0.117)
Treat × Day 16–22	0.093†
	(0.056)
Treat × Day 23–29	0.106†
	(0.059)
Subscriber characteristics	Yes
Time fixed effects	Yes
Observations	150,849
Log likelihood	−51,952

Note. Standard errors are given in parentheses. The omitted category is day −7 to −2.
†$p < 0.1$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

6.3.2. The Postengagement Effect. To investigate subscribers' postengagement behavioral changes, we conduct the dynamic DID analysis presented in Equation (6). Table 6, column (1) reports the regression results using the Poisson DID specification for the full sample. The increase in service consumption is positive and statistically significant in all treatment and posttreatment periods. At the same time, we also observe a decline of the treatment effect over the long run. Specifically, for month 0, the treatment coefficient takes the value of 0.089, which translates to a 9.3% ($=e^{0.089} - 1$) increase in service consumption. For months 1 and 2, we observe a 6.3% ($=e^{0.061} - 1$) increase in service consumption; for months 3 and 4, we see a 5.2% ($=e^{0.051} - 1$) increase in service consumption. These results support Hypothesis 5 (i.e., the persistent posttreatment increase in service consumption may be interpreted as habit formation of subscribers in the treatment group). However, the weakening of such effect over time implies that email engagement, once stopped, failed to induce increased service consumption permanently.

Do treated subscribers differ in terms of the degree to which they form service consumption habits? To answer this question, we conduct three additional heterogeneous analyses to explore postengagement effects among different subscribers. We first conduct the dynamic DID (i.e., Equation (6)) on infrequent and infrequent users of service. Columns (2) and (3) in Table 6 summarize the results of our estimation. The

Table 6. The Postengagement Effect on Service Consumption

	Samples							
	Full (1)	Infrequent users (2)	Frequent users (3)	Basic-level subscribers (4)	Mid-level subscribers (5)	Top-level subscribers (6)	Warm weather (7)	Cold weather (8)
Treat × *DuringTreatment Month*$_0$	0.089***	0.139***	0.101***	0.165***	0.118**	0.143*	0.066	0.102***
	(0.027)	(0.047)	(0.032)	(0.041)	(0.039)	(0.060)	(0.049)	(0.032)
Treat × *PostTreatment Months*$_{1–2}$	0.061**	0.155**	0.053†	0.135***	0.122**	0.068	0.077	0.057†
	(0.027)	(0.048)	(0.029)	(0.038)	(0.037)	(0.064)	(0.052)	(0.030)
Treat × *PostTreatment Months*$_{3–4}$	0.051*	0.125***	0.046	0.137***	0.078*	0.050	0.017	0.063*
	(0.026)	(0.045)	(0.030)	(0.040)	(0.038)	(0.060)	(0.047)	(0.031)
Branch characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Branch fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Subscriber fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	78,231	35,497	42,734	28,442	33,903	15,886	27,206	50,924
Log likelihood	−81,043	−32,477	−48,437	−29,266	−34,614	−17,490	−31,563	−52,668

Notes. Standard errors are given in parentheses. For the heterogeneous analysis on weather, we do not have the location information of 13 subscribers.
$^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

estimated coefficients of treatment are significant for both infrequent users (i.e., 14.9%) and frequent users (i.e., 10.6%) during the treatment month 0. During posttreatment months 1 and 2, the estimated treatment effects are significant: at 16.8% for infrequent users and 5.4% for frequent users. During posttreatment months 3 and 4, the estimated treatment effect is significant (i.e., 13.3%) for infrequent users but insignificant for frequent users at a level of 4.7%. The faster reduction of estimated coefficients for frequent users indicates that the postengagement effect on service consumption is weaker for frequent users than for infrequent users.

Next, we conduct similar regressions for subscribers in different subscription levels. Columns (4)–(6) in Table 6 report the results. Among all three subscription levels, the postengagement effect on service consumption is strongest for basic-level subscribers followed by the mid-level subscribers and weakest for the top-level subscribers. Specifically, for the top-level subscribers, the estimated treatment effects become insignificant immediately after the experiment ended.

Finally, because the treatment in the field experiment took place during the winter season in locations with varying winter climate, we analyze the impact of weather on the engagement effect. To do that, we associate each subscriber with the state that she is in and use the median of the average temperature of these states to classify subscribers into two groups: those living in warm or cold weather states. We conduct the dynamic regression model (i.e., Equation (6)) with respect to each group. Columns (7) and (8) in Table 6 report our estimations. We observe significant treatment effects during all treatment/posttreatment months for subscribers living in cold weather states; however, the treatment/posttreatment effects are statistically insignificant for subscribers living in warm weather states.

7. Data-Driven Email Engagement Strategies

In this section, we seek to answer the following research question. Does email engagement always improve profitability? To answer this question, in Section 7.1, we first conduct a 2×3 (i.e., consumption frequency × subscription level) estimation of the treatment effect of email engagement on subscriber retention and service consumption. Next, in Sections 7.2–7.4, we use the estimation results to conduct a data-driven analysis to optimize the email engagement strategy.

7.1. The Cost-Benefit Trade-Off of Email Engagement

Our paper is motivated by the industry practice of subscription service operations, where customers pay a fixed monthly subscription fee to enjoy oftentimes unlimited service consumption. A subscription service provider's revenue depends on the total number of subscribers, whereas its operating cost depends on all active subscribers' aggregate service consumption. Therefore, there exists an important trade-off in email engagement. That is, when the firm engages its subscribers, its revenue increases because of an increased retention rate; meanwhile, it incurs a higher operating cost because of increased service consumption of retained subscribers. Therefore, email engagement increases both the revenue and the cost, and the net effect on profitability is not immediately apparent without quantifying it from data.

To proceed, we adopt the notion of customer lifetime value, which is defined as the predicted total subscription revenue generated by a customer over her entire life as a subscriber. In other words, we will estimate the long-term effect of email engagement on subscriber retention and service consumption over infinite periods. To do that, we first create six subsamples from the full sample by assigning each customer into one of six groups: whether the customer is an infrequent or frequent user and whether she is enrolled in the basic-level, mid-level, or top-level subscription service. We then conduct regressions to analyze the treatment effect on both retention (i.e., Equations (1)–(3)) and consumption (i.e., Equation (6)) for each group. Our regression estimations are reported in Tables A5 and A6 of the online appendix. We use the estimated consumption increase during the treatment month for numerical calibration, as email engagement is assumed to be repeated on a per-period basis. Technically, to accurately estimate attrition, we need a sufficiently long observation period so that enough customer churns occur. Our study chooses to use five months (entire experiment duration) to estimate the hazard rate, where email engagement stops after one month. Because firms would continuously engage customers with emails in reality, our estimate of the retention benefit of email engagement is conservative.

7.2. The Benefit of Increased Subscriber Retention

To compute customer lifetime value, we first estimate the baseline hazard rate for each customer group. In our sample, the average per-period hazard rates are $h_{ib} = 3.95\%$ for basic-level infrequent users, $h_{fb} = 2.89\%$ for basic-level frequent users, $h_{im} = 4.29\%$ for mid-level infrequent users, $h_{fm} = 3.71\%$ for mid-level frequent users, $h_{it} = 6.94\%$ for top-level infrequent users, and $h_{ft} = 5.49\%$ for top-level frequent users in the control group. Given the estimates of the heterogeneous engagement effects on hazard reduction in Table A4 in the online appendix, the per-period hazard rates with email engagement are $h_{ib}{}' = 3.81\%$, $h_{fb}{}' = 1.94\%$, $h_{im}{}' = 3.87\%$, $h_{fm}{}' = 2.68\%$, $h_{it}{}' = 3.73\%$, and $h_{ft}{}' = 3.22\%$. According to Table 1, the monthly fees paid by basic-level, mid-level, and top-level subscribers are \$14.90, \$24.60, and \$35.70, respectively. These values translate to per-period revenue $R_b = \$7.50$, $R_m = \$12.30$, and $R_t = \$17.90$, respectively. For a customer with a churn hazard rate h and per-period revenue R, we follow the literature (e.g., Fader and Hardie 2007) and calculate the total revenue generated over his lifetime as

$$\sum_{i=0}^{\infty} R(1-h)^i = \frac{R}{h}.$$

Then, the firm's total revenue increase because of increased retention is given by

$$TR = \frac{R}{h'} - \frac{R}{h}.$$

7.3. The Cost of Increased Service Consumption

To estimate the cost of increased service consumption, we communicated with the car wash chain to estimate a set of parameters related to its operating cost. In our context, the operating cost mainly includes electricity cost (\$0.5/wash), natural gas cost (\$0.12/wash), water cost (\$0.16/wash), chemicals cost (\$0.43/wash for basic-level service and \$0.64/wash for mid-level and top-level services), possible repair and maintenance of machinery (\$0.47/wash), and labor and administration cost (\$1.8/wash for basic-level service, \$2.04/wash for mid-level service, and \$2.22 for top-level service). This amounts to a total of $c_b = \$3.48$ for basic-level service, $c_m = \$3.93$ for mid-level service, and $c_t = \$4.11$ for top-level service per wash. These estimates are consistent with survey results reported by the industry newsletter (Auto Laundry News 2016).

Finally, we estimate the average service consumption for each subscription level to calculate the total operations costs. From our data, we calculate that the average per-period consumptions are $q_{ib} = 1.12$, $q_{im} = 1.09$, and $q_{it} = 1.11$ for basic-level, mid-level, and top-level infrequent users, respectively, and $q_{fb} = 2.40$, $q_{fm} = 2.38$, and $q_{ft} = 2.85$ for basic-level, mid-level, and top-level frequent users, respectively, across all periods after the treatment started. According to the results in Table A5 in the online appendix, with treatment, the service consumptions for infrequent users are $q_{ib}{}' = 1.27$, $q_{im}{}' = 1.28$, and $q_{it}{}' = 1.30$; with treatment, the service consumptions for frequent users are $q_{ib}{}' = 2.89$, $q_{im}{}' = 2.67$, and $q_{it}{}' = 3.27$. We then calculate the effect of email engagement on the operating cost for a subscriber's lifetime as follows:

$$TC = c\sum_{i=0}^{\infty} q'(1-h')^i - c\sum_{i=0}^{\infty} q(1-h)^i = c\left(\frac{q'}{h'} - \frac{q}{h}\right),$$

where c is the operations cost of each consumption, q is the consumption without email engagement in period i, q' is the consumption with email engagement in period i, h is the estimated per-period hazard rate without email engagement, and h' is the estimated per-period hazard rate with email engagement.

7.4. Optimizing the Email Engagement Strategy

If we assume the cost of deploying email engagement is negligible (i.e., \$0.0001 as in our collaborator's case), then the net profit of email engagement is given by Profit = TR − TC. With this formula, we can numerically estimate the net profit of email engagement for each subscription level. Table 7 summarizes our estimates.

Table 7. Estimated Financial Impact of Email Engagement

	Basic-level frequent	Basic-level infrequent	Mid-level frequent	Mid-level infrequent	Top-level frequent	Top-level infrequent
Δ profit ($)	−102.3	−10.3	−12.0	1.0	60.0	144.5
95% confidence intervals	[−271.2, −22.3]	[−45.2, 42.4]	[−57.3, 60.4]	[−53.4, 77.7]	[−59.5, 152.4]	[17.9, 325.8]

As it turns out, deploying email engagement on top-level infrequent users is most beneficial, which can lead to a profit improvement of $144.5 for each subscriber. Deploying email engagement on top-level frequent users can lead to a profit improvement of $60. For mid-level infrequent users, the net benefit is much weaker but still positive at $1.0. For mid-level frequent users and all basic-level users, email engagement is counterproductive, as the revenue improvement is offset by a much greater increase of the operating cost because of subscribers' increased service consumption. Strikingly, email engagement on mid-level frequent users and basic-level infrequent users yields net reductions of $12 and $10.3 on profit, respectively. Moreover, basic-level frequent users currently contribute negative profit (=−$29.5) even without email engagement. For this group, conducting email engagement yields a reduction of $72.8 on profit, resulting in a total net profit of −$102.3 per subscriber.

To conclude, we have two managerial recommendations. First, the car wash chain should only target its email engagement program at all top-level subscribers and mid-level subscribers who infrequently utilize service. According to our data, the total fractions of basic-level, mid-level, and top-level subscribers are 30.9%, 40.0%, and 29.1%, respectively. Compared with no email engagement, deploying email engagement on all subscribers would result in a profit improvement of 10.8%, whereas our recommended selective email engagement would result in a profit improvement of 28.6%. Consequently, by adopting our recommendation, the car wash chain can increase its profit by 17.8%. Second, we recommend that the car wash chain adjust its pricing scheme or set a service consumption limit to cut loss on basic-level subscribers.

8. Conclusion

Leveraging a field experiment conducted by a U.S. car wash chain, our study is the first to jointly quantify the causal effect of email engagement on subscriber retention and service consumption in subscription service retailing. We observe that the effect of email engagement on subscriber service consumption exhibits patterns of decay and fatigue, which are consistent with the reminder effect documented in the literature. Furthermore, likely because of subscribers' habit formation, the effect of email engagement persists even after engagement ends, but it weakens over time. Our analysis indicates that email engagement is a double-edged sword that increases both the retention and service consumption of subscribers. From the firm's perspective, a higher retention rate of subscribers increases its revenue; at the same time, additional service consumption increases its operating costs. Therefore, email engagement must be implemented with caution. We use empirical estimations from the field experiment to calibrate a data-driven model to optimize the engagement strategy for heterogeneous subscriber groups. We find that the car wash chain can increase profit by 13.9% if it adopts a selective engagement strategy. Such a strategy can be conveniently implemented at this car wash chain's 130 branches operating in 16 states via easy reprogramming of the CRM software managed by FreshLime, a software as a service company we collaborate with for this project.

More generally, our work is relevant to all subscription businesses where the fulfillment of a physical product or service delivery incurs a substantial marginal operating cost. We hope that our work inspires other companies in the subscription space to reexamine their current email engagement policies and to conduct appropriate cost and benefit analyses. For instance, many online retailers have started to offer subscription box services to their customers (e.g., clothes, jewelry, toy, etc.). Under this business model, a firm sells the product access instead of product ownership to its subscribers for a fixed monthly fee. We note that our paper's findings also apply to the setting of product subscription, where marginal operating costs (i.e., transportation, inventory, and labor costs) are substantial.

Our research demonstrates that combining empirical methods (e.g., field experiments) and personalized data collection technologies (e.g., RFID devices) can enable researchers to investigate interesting consumer behavioral problems in the service sector where activities occur in brick-and-mortar facilities. Personalized data collection devices can allow subscription service providers and retailers to overcome customer tracking barriers and gather granular-level customer data, which open up opportunities for data-driven analytical research.

We shall note that in our experiment, the duration of email engagement is one month, whereas in practice, email engagement can be made indefinitely until customers unsubscribe from the engagement email service. This implies that our field experiment probably captures the lower bound of the effect of email engagement. On the other hand, repeated emails might cause customers to become insensitive to or even annoyed by them and eventually unsubscribe from the email service. Moving forward, a promising future research direction is to implement email engagement experiments that allow different engagement duration and frequencies. It is likely there exists a nonlinear relationship between the effect of email engagement and engagement duration or frequencies. Finding a profit-maximizing email engagement strategy will be an interesting problem to investigate.

Another limitation of our study is that no demographic information is available for the subscribers, as demographic tracking is not easily achieved in brick-and-mortar settings, even with RFID sensors. However, if additional demographic information is available, we can fine-tune the current analysis by segmenting subscribers based on demographic characteristics. This will allow for a more granular analysis and design of email engagement strategies.

Endnotes

[1] In practice, it is common for companies to send engagement emails on a weekly (high frequency), half-monthly (medium frequency), or monthly basis (low frequency). When our collaborator conducted this experiment, they decided to go with the medium frequency of email engagement.

[2] We implemented this three-way fixed effect model using an imported package ppmlhdfe in Stata.

References

Auto Laundry News (2016) Results from the Auto Laundry News: Exterior Conveyor Survey 2016. Accessed June 22, 2022, https://carwashmag.com/wp-content/uploads/ALN_surveys/2016_conveyor_survey.pdf.

Baddeley A (2007) *Working Memory, Thought, and Action* (Oxford University Press, Oxford, UK).

Bai J, So KC, Tang CS, Chen X, Wang H (2019) Coordinating supply and demand on an on-demand service platform with impatient customers. *Manufacturing Service Oper. Management* 21(3):556–570.

Bakos Y, Brynjolfsson E (1999) Bundling information goods: Pricing, profits, and efficiency. *Management Sci.* 45(12):1613–1630.

Belavina E, Girotra K, Kabra A (2017) Online grocery retail: Revenue models and environmental impact. *Management Sci.* 63(6):1781–1799.

Boksem MA, Tops M (2008) Mental fatigue: Costs and benefits. *Brain Res. Rev.* 59(1):125–139.

Cachon GP, Feldman P (2011) Pricing services subject to congestion: Charge per-use fees or sell subscriptions? *Manufacturing Service Oper. Management* 13(2):244–260.

Calzolari G, Nardotto M (2017) Effective reminders. *Management Sci.* 63(9):2915–2932.

Caro F, Gallien J (2010) Inventory management of a fast-fashion retail network. *Oper. Res.* 58(2):257–273.

Caro F, Gallien J (2012) Clearance pricing optimization for a fast-fashion retailer. *Oper. Res.* 60(6):1404–1422.

Charness G, Gneezy U (2009) Incentives to exercise. *Econometrica* 77(3):909–931.

Chen T, Fenyo K, Yang S, Zhang J (2018) Thinking Inside the Subscription Box: New Research on E-commerce Consumers. *McKinsey & Company* (February 9), https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/thinking-inside-the-subscription-box-new-research-on-ecommerce-consumers.

Cui R, Li M, Li Q (2019a) Value of high-quality logistics: Evidence from a clash between SF Express and Alibaba. *Management Sci.* 66(9):3879–3902.

Cui R, Li J, Zhang DJ (2020) Reducing discrimination with reviews in the sharing economy: Evidence from field experiments on Airbnb. *Management Sci.* 66(3):1071–1094.

Cui R, Zhang DJ, Bassamboo A (2019b) Learning from inventory availability information: Evidence from field experiments on Amazon. *Management Sci.* 65(3):1216–1235.

Cui R, Sun T, Lu Z, Golden J (2021) Sooner or later? Promising delivery speed in online retail. Preprint, submitted January 2, http://dx.doi.org/10.2139/ssrn.3563404.

Danaher PJ (2002) Optimal pricing of new subscription services: Analysis of a market experiment. *Marketing Sci.* 21(2):119–138.

Data & Marketing Association (2015) National Client Email Report of U.K. Accessed June 22, 2022, https://dma.org.uk/uploads/ckeditor/National-client-email-2015.pdf.

Du N, Li L, Lu T, Lu X (2020) Prosocial compliance in P2P lending: A natural field experiment. *Management Sci.* 66(1):315–333.

Duflo E (2001) Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *Amer. Econom. Rev.* 91(4):795–813.

Fader PS, Hardie BG (2007) How to project customer retention. *J. Interactive Marketing* 21(1):76–90.

Ferreira KJ, Lee BHA, Simchi-Levi D (2016) Analytics for an online retailer: Demand forecasting and price optimization. *Manufacturing Service Oper. Management* 18(1):69–88.

Fisher M, Gallino S, Li J (2018) Competition-based dynamic pricing in online retailing: A methodology validated with field experiments. *Management Sci.* 64(6):2496–2514.

Ghose A, Li B, Liu S (2019) Mobile targeting using customer trajectory patterns. *Management Sci.* 65(11):5027–5049.

Han R, Chu L, Sun T, Wu L (2022) Commercializing the package flow: Cross-sampling physical products through e-commerce warehouses. Preprint, submitted January 27, http://dx.doi.org/10.2139/ssrn.3566756.

Hu K, Acimovic J, Erize F, Thomas DJ, Van Mieghem JA (2019) Forecasting new product life cycle curves: Practical approach and empirical analysis. *Manufacturing Service Oper. Management* 21(1):66–85.

Hui SK, Inman JJ, Huang Y, Suher J (2013) The effect of in-store travel distance on unplanned spending: Applications to mobile promotion strategies. *J. Marketing* 77(2):1–16.

Karlan D, McConnell M, Mullainathan S, Zinman J (2016) Getting to the top of mind: How reminders increase saving. *Management Sci.* 62(12):3393–3411.

Randhawa RS, Kumar S (2008) Usage restriction and subscription services: Operational benefits with rational users. *Manufacturing Service Oper. Management* 10(3):429–447.

Retana GF, Forman C, Wu DJ (2016) Proactive customer education, customer retention, and demand for technology support: Evidence from a field experiment. *Manufacturing Service Oper. Management* 18(1):34–50.

Rubin DB (1974) Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Ed. Psych.* 66(5):688–701.

Rubin DC, Wenzel AE (1996) One hundred years of forgetting: A quantitative description of retention. *Psych. Rev.* 103(4): 734–760.

Singer JD, Willett JB (2003) *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence* (Oxford University Press, Oxford, United Kingdom).

Staats BR, Dai H, Hofmann D, Milkman KL (2017) motivating process compliance through individual electronic monitoring: An empirical examination of hand hygiene in healthcare. *Management Sci.* 63(5):1563–1585.

SUBTA (2019) 2019 State of the Subscription Commerce Economy Annual Report. Accessed June 22, 2022, https://subta.com/wp-content/uploads/2019/10/2019_SUBTA_Annual_Report.pdf.

U.S. Department of Commerce (2017) New Insights on Retail E-Commerce (ESA Issue Brief #04-17). Accessed June 22, 2022, https://www.commerce.gov/sites/default/files/migrated/reports/new-insights-retail-e-commerce.pdf.

Wooldridge JM (2010) *Econometric Analysis of Cross Section and Panel Data*, 2nd ed. (MIT Press, Cambridge, MA).

Zhang DJ, Allon G, Mieghem JA (2017) Does social interaction improve learning outcomes? Evidence from field experiments on massive open online courses. *Manufacturing Service Oper. Management* 19(3):347–367.

Zhang DJ, Dai H, Dong L, Wu Q, Guo L, Liu X (2019) The value of pop-up stores on retailing platforms: Evidence from a field experiment with Alibaba. *Management Sci.* 65(11):5142–5151.