International Journal of Research in Marketing 38 (2021) 877–899

Contents lists available at ScienceDirect

International Journal of Research in Marketing

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

Full length article

How push messaging impacts consumer spending and reward redemption in store-loyalty programs

Suzanne M.T.A. Bies [a,*], Bart J. Bronnenberg [a], Els Gijsbrechts [a]

[a] Department of Marketing, School of Economics and Management, Tilburg University, Warandelaan 2, 5000LE Tilburg, The Netherlands

ARTICLE INFO

Article history:
First received on January 8, 2020 and was under review for 7 months
Available online 5 March 2021

Area Editor: Peter C. Verhoef

Keywords:
Loyalty programs
Mobile marketing
Field experiment
Heterogeneous treatment effects

ABSTRACT

Retailer loyalty programs (LPs) are pervasive in grocery retailing. However, participant spending and redemption typically wear off over time and traditional communication has not revealed very effective at maintaining program engagement. We study the impact of in-app mobile push notifications on consumer participation and reward collection in store-loyalty programs. Using a unique data set covering consumer spending before and during such a program, we estimate the effect of push messaging on expenditure and reward redemption during the program. We report positive effects of push messages on spending, and even stronger effects on redemption, relative to a control group not receiving such messages. Due to the savings dynamics, the total spending impact is larger for messages sent early on rather than late in the program, while the opposite holds for the total number of stamps redeemed. Conditioning on observable consumer characteristics, we allow for heterogeneous treatment effects and find that the spending and redemption effects of push messaging increase with high levels of pre-program spending. Our findings reveal which loyalty-program stakeholders benefit the most from mobile marketing campaigns, and help to formulate rules for campaign scheduling and targeting.

1. Introduction

Loyalty programs (LPs) are ubiquitous and are in widespread use in the retail-, travel-, and financial services industry, to name a few. These include long-term/permanent LPs, in which points/miles/stamps are saved continuously to be converted into rewards (money or goods), as well as shorter term/temporary LPs that operate for several weeks, after which saved points or stamps lose their value. In 2015, 52% of consumers in the Netherlands participated in such short-term LPs (GfK, 2015). Furthermore, between 2006 and 2015, the use of these LPs by retailers has more than quadrupled in Asia and North America, and even increased six-fold in Europe. The 'market' for such LPs was estimated at about two billion Euros worth of redemption value in 2012 and has been growing since.[1]

In many LPs, consumers receive stamps at the checkout register based on their purchasing amount (e.g., one stamp per 10 euro spent). Consumers can then save these stamps in order to redeem rewards for a pre-specified number at a later point in time, but before the program ends. For short-term LPs, popular reward categories are crystal-, cooking-, and cutting ware, and the rewards can be obtained at a discount or are free. A key component of the success of these programs is 'customer

* Corresponding author.
E-mail addresses: s.m.t.a.bies@tilburguniversity.edu (S.M.T.A. Bies), bart.bronnenberg@tilburguniversity.edu (B.J. Bronnenberg), e.gijsbrechts@tilburguniversity.edu (E. Gijsbrechts).

[1] Based on an internal report of BrandLoyalty, conducted by McKinsey in 2012.

https://doi.org/10.1016/j.ijresmar.2021.02.001
0167-8116/

engagement', i.e. whether the program can activate consumers to save stamps and collect rewards which, in turn, encourages future purchases. Yet, extant studies show that after initial excitement, the salience of, and interest in, LPs typically wears off – as evidenced by declining (purchase and) redemption activity (Dorotic, Verhoef, Fok, & Bijmolt, 2014). Especially light-to-moderate customers of the retailer are found to accumulate points far beyond the redemption thresholds and fail to collect rewards (Kwong, Soman, & Ho, 2011; Stourm, Bradlow, & Fader, 2015; Lal & Bell, 2003; Liu, 2007). Such loss of customer engagement is observed in permanent loyalty programs (PLPs), but also in temporary ones (TLPs), where redemption rates are additionally found to be low (GfK, 2015; Drèze & Hoch, 1998). The critical question is then: How can companies maintain LP salience and consumer participation?

Extant LP studies find little impact of traditional (direct mail or e-mail) communication on program sales or redemption (Lewis, 2004; Dorotic, 2010; Dorotic, et al., 2014). Mobile apps have been advanced as a promising way to increase consumer engagement, in general (van Heerde, Dinner, & Neslin, 2019) and in a LP context in particular (Reinartz & Linzbach, 2018). At the same time, app adoption is not enough: continued app use (stickiness) is important (Wang, Krishnamurthi, & Malthouse, 2018), yet apps, too, "are notorious for a rapid drop in usage" (van Heerde et al., 2019). These authors suggest that repeated push notifications could play a role here, but do not document this empirically (in general or in the context of LPs).

Previous studies on mobile marketing indicate that mobile coupons and advertising text messages can be more effective than direct mail or email (Reichhart, Pescher, & Spann, 2013), and can influence consumers' immediate and planned behavior (Fang, Gu, Luo, & Xu, 2015). However, they also show that the impact of mobile push is context dependent (Bart, Stephen, & Sarvary, 2014) and can, in some instances, be negative (Tsang, Ho, & Liang, 2004). Moreover, also with mobile messages, incrementality is an issue – for instance, with mobile coupons, the question is whether redemption will lead to *additional* sales (Fong, Fang, & Luo, 2015). These issues become particularly relevant for mobile-push reminders in a loyalty-program context, because the repeated messages could come with decreasing returns, and the reminders may trigger consumers to redeem piled-up stamps rather than collect additional stamps (spend more). Therefore, whether and to what extent push notifications increase redemption and sales in the context of a LP (app) remains an unsettled issue.

Our paper intends to fill this gap. Our core objective is to document the impact of mobile push (within apps) on purchases and redemption in a TLP. In so doing, we not only evaluate the effect of individual messages but also consider the impact of the entire push program, including the number and timing of messages.

Additionally, we shed light on the relative impact of push notifications on spending vs. redemption. This is particularly relevant in a TLP setting for two reasons. First, if push messages enhance redemption, this creates a caveat for retailers. On the one hand, redemption may be beneficial because redeemers, feeling more rewarded, are more likely to develop a positive attitude towards the retailer, and to subsequently engage in extra stamp collection (and, thus, spending, see also Dorotic et al., 2014, p. 351). On the other hand, retailers typically incur a cost on each redeemed reward. For retailers, mobile messaging is successful only if it enhances customer engagement in the program while maintaining a sound balance between stamp collection and redemption. Second, short-term LPs are often planned and operated by a third party, the 'program operator,' for profit. The business model of the program operator (and the supplier of reward products) is to make money on each reward redeemed by consumers. So, for these stakeholders, redemption is the primary – if not the only – performance metric.

To determine the causal impact of push messaging, we use a unique panel data set covering a large controlled 18-week field experiment in the Fall of 2016. The data involve 46,504 program-participating consumers shopping at 11,895 stores. Our data track consumers' expenditure and redemption during the short-term LP. In the field experiment, a random subset of panelists (the control group) receives no push messaging after week 3 of the program. The program operator also selects small independent random samples of households who do not receive particular single messages within the program. All other panelists receive the same push messages each week, i.e., without targeting. The program operator varies the number of push messages across weeks from 1 to 3 messages per week.

The random assignment in the communication treatment and the panel structure of our data allow us to use a regression framework that avoids some of the endogeneity concerns in LP- and mobile communication research. For instance, rather than making a contrast between self-selected program participants and non-participants, we compare similar participants exposed to different push messaging regimes assigned at random. In addition, with ample data at the individual level to account for household fixed effects, we can address any unintended selection in the communication experiment (e.g., messages may fail for technological reasons related to household income). For this reason, and being conservative, our preferred specification features two-way (household and time) fixed effects.

Our results show that push messaging strongly enhances the *number of stamps redeemed*, especially among heavy customers at the chain (whose stamp balance quickly builds up). Importantly, though, mobile push also has a strong positive impact on consumer *spending* in the course of the program, and this spending lift, too, increases with high levels of pre-program expenditures. Overall, relative to a random control condition of app users who do not receive push messages after the third week of the program, consumers treated with the push plan redeem about twice as many stamps, and spend about 14% more on average during the program. For the retailer, this results in an 11% margin increase on treated customers throughout the program period – a sizable gain. Hence, unlike traditional communication devices, mobile push messaging clearly enhances program engagement among customers who install the app, and improves performance for the different stakeholders (i.e., the retailer, the program operator, and the manufacturer of rewards). Using our estimates as a basis for simulating the dynamic impact of additional push messages, we uncover managerial guidelines for the targeting and timing of mobile push in short-term LPs. Messages targeted at heavy customers are particularly effective at increasing both stamps

saving and redemption. The timing effect is more ambivalent: back-end loading (i.e. scheduling push late rather than early in the program) increases redemption, whereas messages sent in mid-program weeks have the largest impact on expenditure and stamp collection. Therefore, stamps saving and redemption goals should be carefully balanced when scheduling the push plan.

The remainder of this paper is organized as follows. Section 2 discusses the literature and Section 3 provides the conceptual framework. Section 4 presents the panel data set that tracks program participation and reward collection. Section 5 contains the empirical model, while Section 6 gives the estimation results. Section 7 discusses the implications, and Section 8 concludes.

2. Background Literature

Our paper bridges two streams of literature: that on LPs, and that on mobile marketing. Table 1 summarizes key empirical studies, which we discuss below.

2.1. Loyalty programs

A large body of literature has looked into the effectiveness of LPs (see, e.g., Bijmolt, Dorotic, & Verhoef, 2011; Breugelmans et al., 2015; and Kim, Steinhoff, & Palmatier, 2020 for reviews and discussion). Table 1, Panel A gives an overview of the studies most relevant for our purpose. These field studies shed light on the performance impact of LPs, in consumer packaged goods (CPG) as well as non-CPG settings. While they predominantly focus on permanent LPs, TLPs have received attention as well (Bombaij, Gelper, & Dekimpe, 2020; Drèze & Hoch, 1998; Lal & Bell, 2003; Minnema, Bijmolt, & Non, 2017; Taylor & Neslin, 2005). Overall, these studies indicate that LPs can enhance sales, in a dynamic fashion, and differently so depending on consumers' prior purchase rate at the chain. By awarding stamps or points in proportion to consumers' purchases amounts, LPs – and short-term LPs in particular – may stimulate consumers to spend more at the retailer (e.g. Bijmolt et al., 2011; Breugelmans et al., 2015; Dorotic et al., 2014; Meyer-Waarden, 2007; Lewis, 2004; Taylor & Neslin, 2005; Chauduri, Voorhees & Beck, 2019). These spending effects evolve endogenously. When consumers' stamp balance is still low, the need to accumulate enough stamps to obtain a future reward may motivate them to spend more at the retailer (the 'points-pressure' mechanism, Bijmolt & Verhoef, 2017). Conversely, consumers may also accelerate their purchases as they move closer to their goal of obtaining the reward (the 'goal-gradient hypothesis,' Kivetz, Urminsky, & Zheng, 2006).[2] In all, though, the observed spending increases are often quite modest, especially among consumers with already high pre-program purchase levels at the chain (Liu, 2007; Bijmolt et al., 2011).

Moreover, consumers' attention to, and interest in, the LP is found to decline over time (Dorotic et al., 2014). As the program progresses, consumers – in particular, light-to-moderate buyers – often fail to redeem the collected points (Drèze & Hoch, 1998; Kwong et al., 2011; Stourm et al., 2015), even if the threshold for redemption is exceeded. In some instances, between 50% and 70% of potential LP rewards are never collected (Bijmolt et al., 2011). This loss of engagement is not limited to PLPs: a recent industry report finds that less than half of the participants in TLPs actually redeem a reward (GfK, 2015). To the extent that receiving a reward may revive excitement about the LP (Dorotic et al., 2014) and trigger consumers to step up their expenditures (because of behavioral reinforcement or positive feelings toward the retailer, 'the rewarded-behavior mechanism,' Blattberg, Kim, & Neslin, 2008; Kopalle, Sun, Neslin, Sun, & Swaminathan, 2012), this failure to redeem may be problematic. Continued program participation thus becomes a critical issue.

Even so, extant studies shed little light on how to maintain or increase program salience and engagement. Surprisingly few papers have investigated the impact of retailer communication in the course of the program. Notable exceptions are Lewis (2004), Dorotic (2010) and Dorotic et al. (2014), who study the impact of emails, and Danaher, Sajtos, and Danaher (2020), who include both emails and direct mails.[3] Overall, these studies suggest that retailer communication only has a limited effect on LP sales or redemption.[4] We contribute to this literature by studying the impact of new and possibly more effective forms of direct to consumer communication, in particular (in app) mobile messaging.

[2] These *positive* points-pressure and goal-gradient effects on spending materialize if consumers have not reached the reward threshold yet (that is, when their 'distance' is still positive, and their 'balance' below the minimum number of stamps to collect a reward). In that region, points-pressure increases spending at lower levels of balance (higher levels of distance); and the goal-gradient mechanism lifts sales at higher levels of balance (lower levels of distance). We contend that, as consumers' stamps balance exceeds the reward threshold, (i) the goal-gradient effect is no longer operative (the goal is reached), and (ii) following the points-pressure mechanism, consumers become less inclined to step up their spending as their number of already accumulated stamps grows farther above the redemption threshold.

[3] Wang et al. (2018) include email and direct mail as controls but do not report on the results.

[4] An exception is Danaher, Sajtos, and Danaher (2020), who find that regularly sending direct mails to inform LP members about their point statements encourages program participation. However, their objective and setting is very different from ours. First, they consider a long-term coalition program – with no points expiration, and very little incentive for consumers to step up their purchases at one retailer. Second, every LP member in their setting receives such statements on a regular basis (i.e. there is no 'control group', and no clear time variation in message reception). Though this does not constitute a problem for identifying which types of rewards should be promoted to consumers in different states of LP engagement (their objective), it makes it hard to infer the causal impact of messages. Third, even these authors find that sending emails has little or no influence on program engagement, leaving the question whether other forms of communication (like in-app push) would be effective.

Table 1
Summary of relevant literature.

Panel A: Loyalty Programs

Study	LP		App	Communication		Prior Customer Purchase Rate	Outcome		Dynamics	Setting		Findings
	TLP	PLP		Type	#		Sales-related	Redemption		Real Life	CPG	
Bolton, Kannan &Bramlett (2000)	×	✓	×	×	×	✓	✓[v]	×	✓	✓	×	LP members overlook or discount negative evaluations of company vis-a-vis competition in their repatronage.
Bombaij & Dekimpe (2020)	×	✓	×	×	×	×	✓	×	×	✓	✓	Identify how program design, retailer and country characteristics moderate PLP sales.
Bombaij, Gelper & Dekimpe (2020)	✓	×	×	×	×	×	×	✓	×	✓	✓	Identify how program design, retailer and country characteristics moderate TLP redemptions. Point to difference between PLPs and TLPs.
Danaher, Sajtos, & Danaher (2020)	×	✓	×	DM, email	×	✓	✓	✓	✓	✓	✓	LP members transition between 'activity' states. Promoting specific types of rewards at different states may improve LP effects. DM generally more effective than email.
Dorotic, Verhoef, Fok & Bijmolt (2014)	×	✓	×	DM	×	(✓)	✓	✓	✓	✓	✓	Increasing redemption enhances salience and purchases. Mailings have small impact at best. Impact depends on customers' prior purchases.
Drèze & Hoch (1998)	✓	×	×	×	×	×	✓	✓	✓	✓	✓	TLP increases sales. Low redemption makes it profitable.
Drèze & Nunes (2011)	×	✓	×	×	×	×	✓	×	✓	✓	×	Successive successes (status attainment) enhance effort (extra sales).
Chauduri, Voorhees & Beck (2019)	×	✓	×	×	×	×	✓	×	✓	✓	✓	LPs significantly increase sales and (with a delay) profit in the short and the long run.
Kim, Wang & Malthouse (2015)	×	✓	✓	×	×	✓	✓	×	✓	✓	×	App adoption and continued use increase future spending. When customers discontinue using the app, their spending levels decrease.
Kopalle et al. (2012)	×	✓	×	×	×	✓	✓	✓	✓	✓	×	Both the frequency reward and customer tier components of LPs generate incremental sales.
Lal & Bell (2003)	✓	×	×	×	×	✓	✓	✓	✓	✓	✓	Programs are profitable because incremental sales to casual shoppers offset subsidies to loyal customers.
Leenheer, van Heerde, Bijmolt & Smidts (2007)	×	✓	×	×	×	×	✓	×	×	✓	✓	Small positive effect of PLP membership on share-of-wallet. Correcting for self-selection = important.
Lewis (2004)	(✓)*	(✓)*	×	email**	×	✓	✓	×	✓	✓	✓	Accounts for long term effects. LP successful for increasing annual sales.
Liu (2007)	×	✓	×	×	×	✓	✓	✓	✓	✓	✓	Heavy customers before the LP redeem more but do not buy more, light customers buy more and become more loyal.
Liu & Yang (2009)	×	✓	×	×	×	✓	✓	×	×	✓	×	With competing LPs. larger firms benefit more from their LPs. In non-expandable categories, LP effect erodes with # of programs.
Mägi (2003)	×	✓	×	×	×	×	✓	×	×	✓	✓	Impact of loyalty cards on consumer share-of-wallet is mixed. Competing cards may cancel each other out.
Meyer-Waarden (2007)	×	✓	×	×	×	✓	✓	×	✓	✓	✓	LPs increase customer lifetimes. The effect decreases with multiple card memberships of nearby stores and is higher for high-share customers.
Meyer-Waarden & Benavent (2009)	×	✓	×	×	×	✓	✓	×	✓	✓	✓	Heavier (lighter) buyers adopt loyalty cards earlier (later). Late card adopters do not buy more at the chain in the longer run.
Minnema, Bijmolt & Non (2017)	✓	×	×	×	×	✓	✓	×	✓	✓	✓	IRP results in incremental shopping trips. Bonus premiums enhance the impact of price cuts on brand sales. IRPs and bonus premiums are especially (but not exclusively) effective for households that collect the premiums.
Rossi (2018)	×	✓	×	×	×	✓	✓	×	✓	✓	×	LP effects are driven by a segment of frequent customers who value rewards more than money (price-insensitive reward-seeking = PIRS consumers).
Sharp & Sharp (1997)	×	✓	×	×	×	✓	✓	×	✓	✓	✓	LPs do not systematically increase sales at participating stores.
Stourm, Bradlow & Fader (2015)	×	✓	×	×	×	×	×	✓	×	✓	×	Redemption is low. Customers differ in their motivation to redeem. Redemption behavior is driven by cognitive and psychological incentives.
Taylor & Neslin (2005)	✓	×	×	×	×	×	✓	✓	✓	✓	✓	TLP = profitable. Find both a points-pressure and rewarded-behavior impact. TLP can appeal to customers not interested in the PLP.
Wang, Krishnamurthi & Malthouse (2018)	×	✓	✓	DM, email	×	✓	✓	✓	✓	✓	✓	Mobile app and a flexible reward structure increase sales and redemption from LPs (especially from occasional customers), and spending at partners that sell high-penetration categories.
Wang, Lewis, Cryder & Sprigg (2016)	✓	×	×	×	×	×	✓	×	✓	✓	×	TLP increased purchasing in the post-promotion period. Marketers should set reachable goals, and be cautious about goal failures among loyal customers.
Zhang & Breugelmans (2012)	×	✓	×***	×	×	×[iv]	✓	×	×	✓	✓	IBLPs increase (decrease) sales among previous non-LP (LP) members, reduced attrition, and attracted more new customers.

DM = direct mail, # = number of messages. *The program has an annual time constraint, **email coupon independent of LP, *** the retailer used emails and store flyer announcements, but these could not be distinguished from the program change itself, [iv] prior LP membership accounted for, not sales, [v] repatronage

Panel B: Mobile Marketing: Apps and Mobile Communication

Study	LP		App	Communication		Heterogeneity: Prior Purchases	Outcome		Consumer Response Dynamics	Setting		Findings
	TLP	PLP		Type	#		Sales	Redemption		Real Life	CPG	
Mobile Apps												
Gu & Kannan (2021)	×	×	✓	email	×	✓	✓	×	✓	✓	×	In competitive setting (hotels), app adoption decreases spending, especially when used for search and with low app engagement.
Kim, Wang & Malthouse (2015)	×	✓	✓	×	×	✓	✓	×	✓	✓	×	Sticky apps can lift brand sales, poorly designed apps can hurt the brand.
Li, Luo, Zhang & Wang (2017)	×	×	✓	m-ad, SMS & in-app	×	×	✓	×	✓	✓	×	Consumer response to mobile promos depends on weather. Prevention ads lower (improve) the promo impact in sunny (rainy) weather.
Narang & Shankar (2019)	×	×	✓	×	×	✓	✓	×	✓	✓	×	App adopters spend more, return more, and yield higher net sales, online and offline. Sales depend on app usage rather than availability.
Van Heerde, Dinner & Neslin (2019)	×	×	✓	×	×	✓	✓	×	✓	✓	×	Sales increases from retailer apps are higher for distant and for offline-only customers, due to higher app access (use).
Zubcsek, Katona & Sarvary (2017)	×	×	✓	m-coupon, in app	×	✓	×	✓	×	✓	✓	Co-located consumers respond to coupons in the same product category. Location history can be used to target next-day coupons.
Study	LP		App	Communication		Heterogeneity: Prior Purchases	Outcome		Consumer Response Dynamics	Setting		Findings
	TLP	PLP		Type	#		Sales	Redemption		Real Life	CPG	
Mobile Communication												
Andrews, Luo, Fang, & Ghose (2016)	×	×	×	m-ad, SMS	×	×	✓	×	×	✓	×	Physical crowdedness increases consumer response to mobile ads.
Bart et al. (2014)	×	×	×	m-ad, webpage	×	×	(✓)*	×	✓	✓	✓	Mobile ads are often ineffective, but can enhance recall and purchase intentions for high-involvement and utilitarian products.
Danaher et al. (2015)	×	×	×	m-coupon, SMS	×	×	×	✓	✓	✓	✓	Impact of m-coupons depends on face value, delivery time and location. Work best for short expiry times.
Dubé, Fang, Fong & Luo (2017)	×	×	×	m-coupon, SMS	×	✓	×	✓	×	✓	×	Competitive responses raise (lower) the profitability of behavioral (geographic) targeting; (a) symmetric pricing incentives soften (toughen) price competition.
Fang et al. (2015)	×	×	×	m-ad, SMS	×	×	✓	×	✓	✓	×	Location-based mobile promotions (LMPs) induce impulsive same-day purchases, but also create product awareness and sales for subsequent days (12). Location matters.
Fong, Fang & Luo (2015)	×	×	×	m-coupon, SMS	×	×	✓	×	×	✓	×	Promoting to consumers near a competitor's location generates extra sales without cannibalizing profits.
Ghose, Li & Liu (2019)	×	×	×	m-coupon, SMS	×	×	✓	✓	✓	✓	✓	Trajectory-based mobile targeting within a shopping mall increases redemption and (current and future) spending, especially for high-income consumers, during weekdays and for exploratory shopping.
Liu, Lobschat, Verhoef & Zhang (2019)	×	×	✓	×	×	✓	✓	×	✓	✓	×	Adding an app to a retailer's mobile website can increase customers' purchase incidence, frequency and order size, and enhance loyalty.
Luo, Andrews, Fang & Phang (2014)	×	×	×	m-ad, SMS	×	×	✓	×	✓	✓	×	Promoting at close (far) distance is more effective with same-day (one-day prior) targeting.
Mills & Zamudio (2018)	×	×	×	m-coupon, self-scan	✓	✓	×	✓	✓	✓	✓	Sheds light on drivers of m-coupon redemption under competition (coupon value, price range, and brand loyalty).
Osinga, Zevenbergen & van Zuijlen (2019)	×	×	×	m-banner ad	×	×	✓	×	✓	✓	×	Mobile banner ads increase offline sales, but do not affect online sales.
Park, Park & Schweidel (2018)	×	×	×	m-coupon, SMS	✓	✓	✓	×	✓	✓	×	Price discount and free sample mobile coupons increase consumers' purchase likelihood and spending during the redemption period. For free sample coupons, the effect lasts beyond that period.
Wang, Krishnamurthi & Malthouse (2018)	×	✓	✓	DM, email	×	✓	✓	✓	✓	✓	✓	Mobile app and a flexible reward structure increase sales and redemption from LPs (especially from occasional customers), and spending at partners that sell high-penetration categories.
Our Study	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	

#Number of messages. *Purchase intention. Note: Table1 includes empirical (field) studies, as published in JM, JMR, Mkt Sc, Man Sc, JAMS, IJRM and JR.

2.2. Mobile marketing

A second stream of literature to which we aim to contribute is mobile marketing, e.g., the use of mobile apps, and mobile communication (see e.g. Andrews, Goehring, Hui, Pancras, & Thornswood, 2016; Grewal, Bart, Spann, & Zubcsek, 2016; Lamberton & Stephen, 2016; and Shankar et al., 2016 for reviews and discussion). Table 1, Panel B summarizes key field studies on the performance impact of mobile marketing.

Mobile Apps. Mobile applications ('apps') have been advanced as a potential way to engage consumers and stimulate spending, in general (Grewal et al., 2016, Shankar et al., 2016, van Heerde, et al., 2019), and in the context of LPs: "Mobile becomes a highly viable channel for customized application-based [...] reward programs" (Reinartz & Linzbach, 2018, p. 316). Several studies empirically report on the impact of retailer apps in general, and show that they can increase online and offline sales (e.g. Liu et al., 2019; Narang & Shankar, 2019; van Heerde et al., 2019; Zubcsek, Katona, & Sarvary, 2017). Still, Gu & Kannan (2021) find that in a highly competitive setting (i.c., hotels), app adoption can reduce spending, especially among consumers with low app engagement. In a loyalty-program setting, Wang et al. (2018) find that linking a retailer's LP to a mobile app may increase program sales and redemption. However, Kim, Wang and Malthouse (2015) show that when consumers discontinue using the app, their spending levels decrease again. Sustained sales lift thus only comes with continued app usage ('sticky apps', Kim et al., 2015), yet consumers are found to quickly lose interest in apps over time (Appel, Libai, Muller, & Shachar, 2020; van Heerde et al., 2019). Repeated notifications sent from within the app might act as reminders and maintain interest, but this is not documented yet empirically.

Mobile communication. Besides apps, mobile push communication, that is, "content sent by or on behalf of advertisers and marketers to a mobile device at a time other than when the subscriber requests it" (Unni & Harmon, 2007, p.30) has also received extensive attention in the academic literature (see Table 1, Panel B). This form of direct communication has gained appeal with the growth of smartphone use (Grewal et al., 2016). Extant studies on mobile communication have mainly focused on 'stand-alone' mobile coupons (Danaher, Smith, Ranasinghe, & Danaher, 2015; Dubé, Fang, Fong, & Luo, 2017; Park, Park, & Schweidel, 2018; Zubcsek et al., 2017; Mills & Zamudio, 2018; Fong et al., 2015) or mobile ads (Andrews et al., 2016; Bart et al., 2014; Luo, Andrews, Fang, & Phang, 2014; Fang et al., 2015; Osinga, Zevenbergen, & van Zuijlen, 2019; Li, Luo, Zhang, & Wang, 2017), mostly delivered as text messages. Apart from reaching consumers in real-time and at low cost, these mobile (push) messages are found to grab attention, trigger recall, provide contextually-relevant content, and foster intimacy/engagement (Bellman, Potter, Treleaven-Hassard, Robinson, & Varan, 2011; Grewal et al., 2016; Bart et al., 2014; Shankar et al., 2016; Reinartz & Linzbach, 2018) – thereby increasing sales of the promoted items in the immediate period and beyond (Fang et al., 2015). However, push messaging also has downsides. (Too many) messages may lose effect and even lead to irritation (see, e.g., Tsang et al., 2004). Moreover, being sent to personal devices, mobile push messages can be perceived as intrusive and cause privacy concerns (Andrews et al., 2016), which, in turn, may negatively affect purchases behavior (Phelps, D'Souza, & Nowak, 2001). Finally, for push messages that offer a benefit to the consumer, incrementality can be an issue – the question being whether the additional sales make up for the cost of granting the benefit (Fong et al., 2015).

While extant papers (as shown in Table 1, Panel B) have documented the impact of (i) 'stand-alone' mobile messages (coupons or ads) (ii) for individual products (iii) typically distributed through SMS, no study to date has studied the impact of (i) push-notification schedules, (ii) in the context of LPs, (iii) sent within a mobile application.[5] Yet, the effect of the latter messages may be quite different. On the one hand, push messages sent within a mobile application may receive more attention and suffer less from privacy concerns, because consumers opted in for the app. Moreover, LP-related notifications are typically not limited to a particular promoted item but pertain to the entire spending amount at the retailer, and, to the extent that these messages alert consumers to already 'earned' rewards, they may generate a more positive response. On the other hand, the use of *repeated* messages within the LP may lead to a loss in effectiveness, and it remains unclear whether the incremental sales for the retailer – if any – make up for the cost of increased redemption.

To summarize this section, extant studies on LPs highlight the need to enhance and maintain program engagement (salience and redemption), but fail to identify forms of communication that clearly serve this purpose. In turn, the literature on mobile marketing emphasizes the potential of mobile apps and messages to trigger attention and foster sales, but does not document to what extent these effects apply to in-app push-notification schedules for LPs. Our paper intends to fill this gap.

3. Conceptual framework

3.1. Overall framework and setting

Fig. 1 presents our conceptual framework. Our two outcomes of interest are consumer spending (and, hence, *number of stamps earned*), and redemption (the *number of stamps traded in for rewards)* during the program. Our framework accounts for several shifters of savings and redemption (dotted lines) that operate independently from push messaging. These are consumers' *stamp balance*, *distance to the reward threshold*, and *previous redemption*. Our focus, though, is on how mobile push

[5] Li et al. (2017) study in-app push messages, but focus on individual messages outside of a LP context.

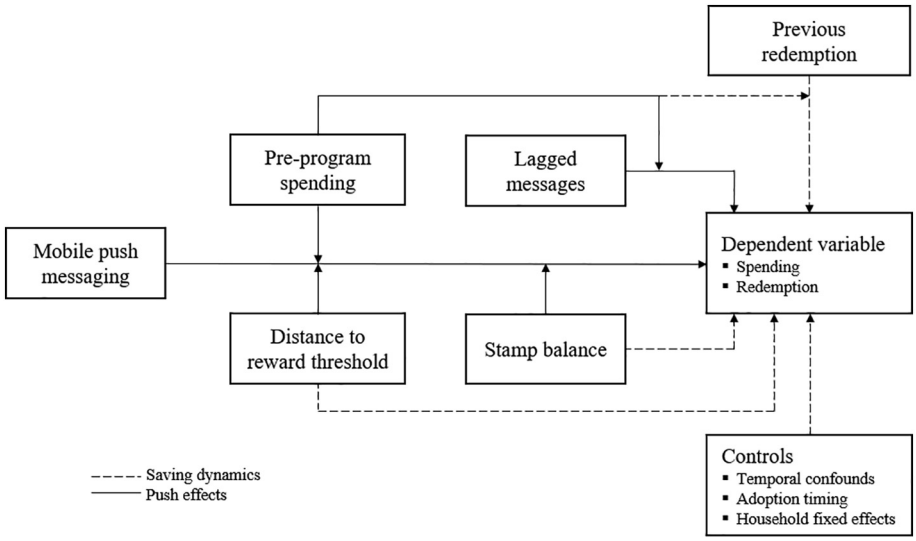

Fig. 1. Conceptual framework

messages (left side of Fig. 1) intervene in these processes and ultimately affect in-store spending and stamp redemption (solid lines in Fig. 1). We discuss our expectations about these effects below.

3.2. Immediate effects

Main message effects. As indicated above, consumer engagement in LPs typically declines over time, because the program becomes less salient and/or consumers lose interest (Dorotic et al., 2014). Push messages can help to remedy this. First, they can serve as a simple reminder of the program and the associated rewards. Second, mobile messages can strengthen latent consumer goals, including reward-seeking (Shankar et al., 2016) and thus help increase consumers' stamp saving efforts and in-store spending. Third, push messages may remind consumers of the efforts already put in thus far (e.g. extra trips or purchases at the retailer), which will become sunk should the consumer not keep up those efforts or fail to see them through by redeeming a reward ('switching costs', e.g. Kim, Shi, & Srinivasan, 2001; Taylor & Neslin, 2005; Kivetz, 2003). On the downside, as more messages are sent, these messages may come to lose effect (and, possibly, even cause irritation). Moreover, being sent to personal devices, mobile messages may create privacy concerns. However, to the extent that push notifications are only sent to consumers who voluntarily opted in for the app, we do not expect the latter to be too much of a problem. So, on the whole, we expect a positive but marginally diminishing main effect of messages on spending and redemption.

Interactions with distance and balance. Building on Kivetz et al.'s (2006) observation that goal proximity increases promotion sensitivity, we anticipate higher push effects when distance is low, as the consumer is reminded that he put in a lot of effort already and is almost there. We expect a negative interaction effect on spending between push-messages and balance. On the one hand, when balance is still (well) below the reward threshold, the message may remind consumers of what they still have to accomplish, and provide a trigger to save for a reward (Koo & Fishbach, 2008). At the same time, as consumers' stamp balance grows beyond the threshold, messages may make consumers more aware not only of the program, but also of the already collected stamp surplus, and become less effective in stimulating extra spending. As for redemption, we propose a larger message effect when balance is high and consumers can collect their reward without (much) further effort. Conversely, when the distance is too high, the consumer still cannot redeem any stamps even if pushed, so we expect a negative interaction on redemption.

3.3. Dynamics and Heterogeneity

Lagged effects. Besides an immediate impact, we also expect push messages to have a 'lagged' effect, i.e. affect consumers' spending and redemption behavior in the next period(s). For spending, this effect can go both ways. On the one hand, it may be positive due to a delayed message-response (e.g. Braun & Moe, 2013): consumers may only (have the opportunity to) visit the chain in the period(s) following the message and, being reminded of the program, may increase their expenditures at that time. On the other hand, similar to a post-promotion dip (see, e.g., Neslin & van Heerde, 2008), the lagged effect may be negative due to purchase acceleration: consumers who increased their purchases at the time of the message and built up inventory may subsequently buy less. Because it is not clear upfront which of these two forces dominate, we have no directional expectation for the lagged-message effect on spending. As for redemption: ceteris paribus (that is: controlling for stamp balance), we do not expect previous-period messages to drive down the current number of stamps redeemed. Rather, these

push messages may have made the reward program more salient and therefore serve as a reminder to trade in stamps for rewards.

Pre-program spending. As discussed above, consumers' pre-program purchase levels at the retailer affect their program participation (spending and redemption). An important question, then, is how it shapes their response to mobile push messages. For spending, the answer could go either way. On the one hand, lighter buyers may find the redemption-threshold more difficult to attain, and be less prone to spend more in response to a push message. (Too many) push messages targeted at this segment could even produce a reactance effect and lead consumers to spend less (Stauss, Schmidt, & Schoeler, 2005). Moreover, these buyers may not be in the habit of visiting the chain on a frequent basis, and thus have less opportunity to (immediately) step up their expenditures unless the message convinces them to incur an extra visit. On the other hand, while heavier buyers who regularly visit the chain may have more *opportunity*, they may also have less *room* for further spending increases – the so-called ceiling effect (Bijmolt et al., 2011). We leave the net outcome of these forces on consumers' immediate and lagged message response as an empirical issue. When it comes to redemption, we expect higher pre-program purchase levels to enhance the push-messaging effects: heavier spenders can more easily surpass the redemption threshold whenever they receive a message and trade in their stamps for a reward in response, while low-spenders may become frustrated by these messages (Stauss et al., 2005).

Given that redemption not only incites future spending but also comes at a cost, a key question for the retailer is how the spending and redemption impact of push messaging net out, over time, for each customer group – something our estimation results will shed light on.

4. Data

4.1. Program details and data sources

Our data cover an 18-week LP, operated by BrandLoyalty, and running from August 1, 2016 until November 30, 2016, across 12,135 stores in a major retail chain in Indonesia. Consumers save for rewards by collecting stamps, earned at a fixed rate of one stamp per 40,000 IDR (Indonesian Rupiah) – equivalent to approximately US$2.8 – spent in one of the chain's stores. Stamps can be collected physically or digitally. Digital collection requires installing an app, available from the start of the program onwards. Throughout the program, push messages can be sent to the shoppers with this app. While installing the app is not required to participate in the program, consumers only start receiving push messages after they install the app. To use the app, consumers need to link the installed app to their customer card. When a customer card is scanned at the check-out, the consumer's stamp balance is automatically updated by the number of stamps earned or redeemed.

The rewards for this program are plastic containers to store food, offered at a large discount. To obtain a reward, consumers hand in a pre-specified number of stamps, and pay a strongly reduced price (below the wholesale price of the reward for the retailer). The containers vary in size and in the number of stamps required to redeem each item. Web Appendix 1 provides details on the reward items and associated stamps requirements. Consumers can collect and redeem stamps during the first 14 weeks of the program but can only redeem stamps during the final 4 weeks (clean-up weeks). However, during these final weeks, the program operator only pushes messages to consumers who have a balance high enough to still redeem a reward. To avoid potential bias, we estimate the effects based on the 14 program weeks only, and we exclude these final clean-up weeks. Note that all treated consumers receive the same number and type of push messages during program weeks, i.e., the program operator does not target specific groups of consumers during the program weeks.

Because our objective is to assess the impact of push messages sent within the program app, our population of interest consists of consumers who adopt the app (and hence, by necessity, are also card holders of the chain). As part of the measurement system to evaluate the push campaign, the program operator designed exogenous variation in push messages in two different ways. First, a randomly selected subgroup of 2305 app-installers only receive push messages in the first 3 weeks of the program, and no messages thereafter. These customers serve as the 'control group.' The construction of this group selects on app installation in the first 2 weeks of the program. The remaining customers, who also opted in for the app, do receive push messages in later program weeks, and make up the 'treatment group.' We focus on those households in the treatment group who install the app in the exact same period as the control households, i.e., the first 2 weeks of the program. This ensures that the control and treatment group are comparable on unobservables that correlate with the app installation timing, and leaves us with 44.2K treated shoppers who received push messages throughout the program.

In addition, the program operator withheld individual push messages from a subsample of the consumers in the treatment condition, throughout the program. This resulted in randomly selected subsamples that received one message less in a given week. Depending on the week, the size of such a subsample was 3-9% of the 44.2K shoppers. Each subsample was drawn independently across messages from the pool of treated consumers. Membership of these additional control groups is not correlated with pre-program spending levels ($\rho = .001, N = 651,056, p = 0.358$). The resulting variation in push messages is exogenous. Finally, the delivery of messages through the mobile app may fail for additional reasons including canceled phone service, outages, and the like. The fraction of failed messages is not trivial: of the 651K shopper-weeks

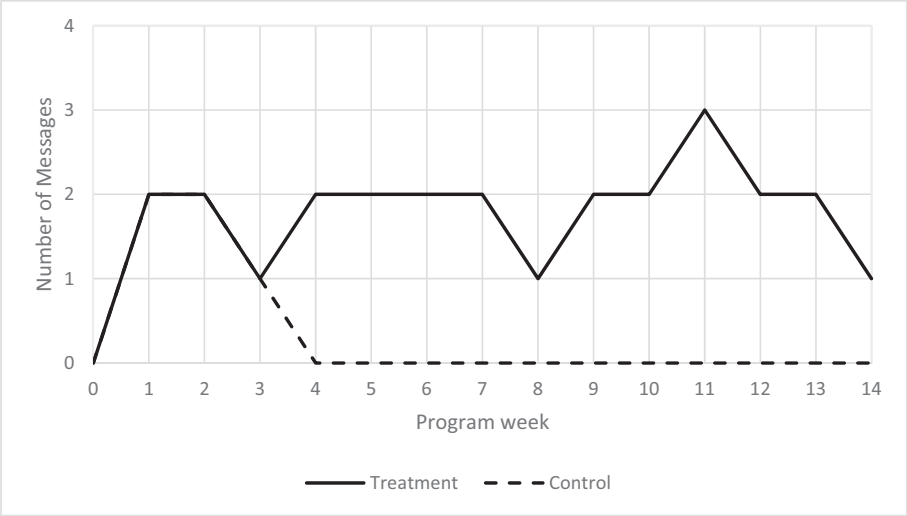

Fig. 2. Number of push messages sent per week of the program

with messages, 26% have at least one failed message, and message delivery to a given customer can be persistently poor. However, these failures are not the result of targeting by the retailer or program operator. Also, the correlation between missing messages and spending levels prior to the program is small (albeit significant given the large number of observations; $\rho = -.038, N = 651,056, p < 0.001$).[6] As such, the failed messages provide an additional source of variation in the number of messages received. The push message schedule as used by the program operator is shown in Fig. 2.[7]

With these program details in mind, our study merges two datasets. The first dataset contains app information at the household-day level, and captures when consumers go the store, how many stamps they redeem on each transaction, and when they receive push messages. The app data from this first dataset only provides information on the number of stamps redeemed from the point a consumer adopts the app until the end of the program.[8] The second, complementary, dataset contains the exact spending amounts for the same households, obtained from the chain's database, for 18 weeks before the start of the program and for an additional 18 weeks covering the program. To combine these two sources, we first aggregate messages, transactions, and stamp redemption information from the app data to the household-week level.[9] The merged dataset then contains information for households that subscribe to the app at some point during the first 2 weeks of the program. The data also cover households' purchase history with the retailer prior to the program. We will later use this information to distinguish between more-vs.-less heavy buyers at the chain. Because we are interested in the effect of push messages and not the program itself, the analysis of push message effectiveness uses the periods covering the program (during which push messages could be sent). In total, our data cover 651K household-weeks for estimation.

4.2. Sample statistics

Table 2 provides sample statistics on the average weekly expenditures and number of visits to the chain running the LP. It does so for the sample as a whole, and also separately for the treatment and control group. In addition, it shows the split between subgroups with different levels of pre-program spending: the 'low-quartile' (lowest-25% customers), 'median' (25–75%), and 'high-quartile' group (top-25% spenders).

During pre-program weeks, consumers spend 167,400 IDR per week, and undertake 2.33 weekly trips to the chain, on average. At the same time, we observe large variation within the customer base: mean spending amounts and visits prior

[6] We note that even if there is a link between number of failed messages and customer spending prior to the program, this does not jeopardize our identification of the message effects because our model includes household fixed effects and interactions with pre-program spending.

[7] We are not aware of the options to block push messages, nor do our data offer information in this regard. However, even if consumers can block push messages and some did, the results from our field experiment still show that push messaging has a positive impact on spending and redemption.

[8] Though we do not observe redemption behavior before app adoption, we do not believe this renders our study invalid for two reasons. First, given the observed redemption behavior and pace of accumulating stamps, there will generally be low incidence of redemption in the first two weeks. Second, such redemption cannot be the result of push messages because the consumer receives none. In addition, we account for household specific fixed effects, so heavy users' baseline redemption is taken into account in our analysis.

[9] To the best of our knowledge, messages are sent to everyone who is a recipient at the same time. There is likely variation with respect to when households read the message even when it is sent at the same time but we have no information on this. We only know when messages are received, not whether or when they are read. If a message is received in the evening, it may not be read until the next morning. Aggregation to the week level helps overcome this issue, and we assume that this variation does not impact our weekly results.

Table 2
Household shopping activity by treatment group, program status and household quartiles.

	Mean	SD	N	0–25% HH Mean	25–75% HH Mean	75–100% HH Mean
Weekly expenditures (1000 IDR)						
Pre-program						
Control	163.4	218.1	2,305	20.3	91.6	455.9
Treatment	167.6	225.6	44,199	20.4	93.1	463.5
Total	167.4	225.3	46,504	20.4	93.1	463.2
During program						
Control	155.5	184.3	2,305	59.9	117.5	331.4
Treatment	169.9	199.5	44,199	62.6	129.4	358.0
Total	169.2	198.8	46,504	62.5	128.8	356.7
Number of trips						
Pre-program						
Control	2.28	2.64	2,305	0.57	1.57	5.49
Treatment	2.33	2.64	44,199	0.57	1.59	5.57
Total	2.33	2.64	46,504	0.57	1.58	5.57
During program						
Control	2.31	2.31	2,305	1.04	1.90	4.47
Treatment	2.48	2.47	44,199	1.11	2.03	4.74
Total	2.47	2.47	46,504	1.10	2.03	4.73

to the program range between 20,300 IDR (.57 visits) for the low-quartile group, and 463,500 IDR (5.57 visits) for the high-quartile customers. This will make the analysis of differences in push-message responsiveness between lighter and heavier buyers particularly relevant. Comparing consumers' behavior before and during the program, we observe a slight increase in weekly spending (on average: from 167,400 to 169,200 IDR) and visit frequency (on average: from 2.33 to 2.47 visits per week). These increases cannot be interpreted as a main effect of the loyalty program, because they are confounded with a time trend among program participants.

Our interest, though, is not with the main effect of the LP, but with the impact of push messaging on consumer saving and redemption behavior, which we will assess by monitoring consumers in the control versus the treatment group. Comparing the pre-program behavior of these two groups, we find that pre-program spending levels are comparable (163,400 IDR for the control group vs. 167,600 IDR for the treatment group), as are the store visits (2.28 vs. 2.33 times per week). We conducted a parallel-trends test in the pre-program periods by testing for the equality of time trends across the treatment vs. control groups. Applying a linear trend model to the pre-program expenditure data, the difference in trends between the treated and control group is very small at 100 IDR per week (s.e. = 394.87) and we cannot reject the null hypothesis that both groups exhibit similar purchase dynamics prior to the program. Fig. 3 shows the average weekly household spending for the treated and control households across different periods in time. It confirms that the expenditures of the treated and the control are highly similar in the pre-program weeks.

Zooming in on the shifts from pre-program to within-program weeks in Table 2, we observe that the control group spends about 155,500 IDR during program weeks (against 163,400 IDR pre-program). As above, this difference cannot be interpreted as a program effect. The effect of push messaging is informed by comparing this drop to the spending shift for the treatment group. For the treated households who keep receiving push messages, expenditure is almost the same during the program (169,900 IDR) as before (167,600 IDR). This is also visible in Fig. 3. Although the pre-program trends are similar, from week 34 onward, as the control households no longer obtain messages, the participants who obtain push messages spend consistently more on average than those who do not obtain push messages. Similarly, whereas households in the control group display almost the same shopping frequency during the program as they did before (2.31 vs 2.28 trips per week), households in the treated group show an increase (from 2.33 to 2.48 trips per week). These statistics give a first indication that push messaging enhances the outcomes of the program – an effect that we will separate out more carefully with our formal modelling approach.

Table 3 provides more detailed statistics on households' saving behavior during program weeks. It reports sample statistics of stamp collection, fraction of weeks with redemption, and number of stamps redeemed per week, in the treatment and control group and for households with different pre-program spending levels. Several points are worth noting. First, stamp collection and redemption monotonically increase with consumers' spending prior to the program. Second, with a mean of 4.08 stamps collected per week, the average consumer should not find it too hard to exceed the modal redemption threshold of 10 stamps (which we will use as the threshold in our main models).[10] However, as we already saw for expenditures, the

[10] We also considered a threshold of five stamps (the lowest requirement, which applied to one of the rewards). As can be seen from Web Appendix 5, the pattern of results remained the same. Because the 5-stamps reward was not a collectable (unlike the other rewards), and because the majority of rewards required (at least) 10 stamps (see Web Appendix 1), we retained the 10-stamps threshold in our main model.

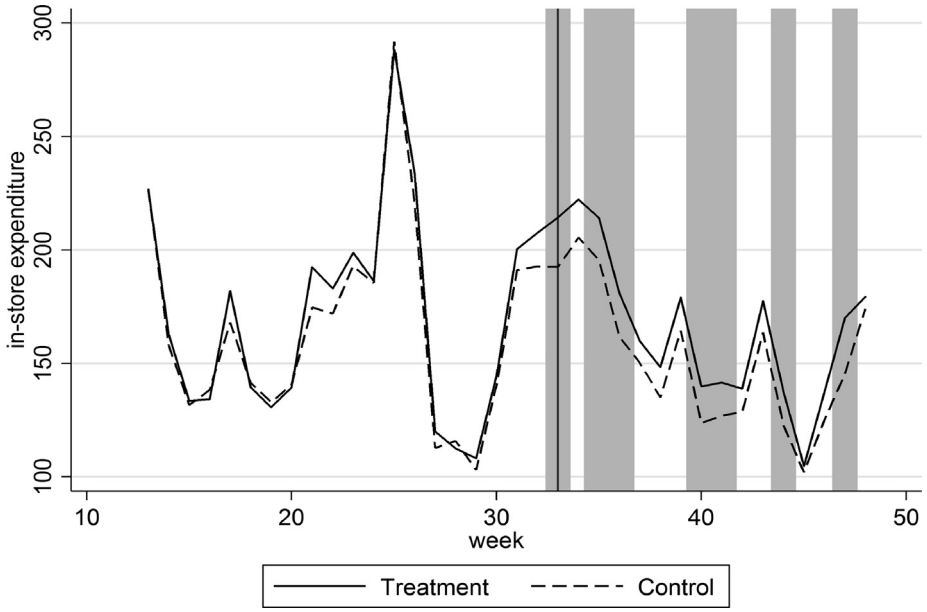

Fig. 3. In-store expenditure by treatment and time. Note: The shaded time-periods denote weeks with significant differences in the average in-store expenditure between households in the treated and the control groups. The vertical line denotes week 33 after which differences in treatment start. None of the differences between the treated and the control group are significant before week 33.

Table 3
Heterogeneity in program participation and rewards redemption.

	Weekly stamp collection		Fraction of redemption incidence		Stamps redeemed per week		N
	Mean	SD	Mean	SD	Mean	SD	
Control							
Average household	3.73	4.80	0.13	0.16	2.91	4.50	2,305
0–25% households	1.33	2.89	0.06	0.11	1.04	2.81	598
25–75% households	2.72	2.89	0.12	0.13	2.09	2.97	1,136
75–100% households	8.25	6.25	0.26	0.19	6.52	6.15	571
Treatment							
Average household	4.10	5.18	0.15	0.17	3.25	4.85	44,199
0–25% households	1.42	2.69	0.06	0.10	1.09	2.61	11,028
25–75% households	3.01	3.34	0.12	0.13	2.31	3.28	22,116
75–100% households	8.93	6.70	0.28	0.20	7.27	6.56	11,055
Total							
Average household	4.08	5.16	0.15	0.17	3.23	4.83	46,504
0–25% households	1.42	2.70	0.06	0.10	1.09	2.62	11,626
25–75% households	3.00	3.32	0.12	0.13	2.30	3.26	23,252
75–100% households	8.90	6.68	0.28	0.20	7.23	6.54	11,626

distribution of weekly stamp collection across consumers is heavily skewed, with a value of only 3.01 for the median group, and only 1.42 for the low-quartile group – indicating that some consumers have to save for weeks before they can collect a reward. Third, redemption is low. For instance, with 3.00 stamps collected per week, a 'median' household accumulates 42 stamps in the course of the 14 program weeks. However, the number of stamps actually traded in for a reward is much lower, at about 32 stamps (=14 $\times$ 2.30). The same holds for heavy buyers, who accumulate some 125 stamps in the course of the program, only 101 of which are redeemed. So, even consumers who expressed an interest in the program by installing the app often fail to redeem their rewards. This underscores the importance of mobile push messaging as a potential tool to increase redemption.

5. Model

5.1. Empirical issues

Recall from Section 4.1 that our estimation sample includes households who, like the control households, installed the app in the first two weeks. This allows us to study the effect of push messaging holding installation timing of the app constant and to rely solely on variation from random assignment into treatment and control.

In our empirical modeling, we further need to account for a number of confounds. First, in addition to a program effect, other temporal factors (e.g. trends, holidays) can influence households' stamp saving or redemption behavior. We control for these factors through time fixed effects. Second, the adoption of the app in itself may lead to increases in stamp saving. To separate this app effect from the message effect, we include a step dummy that indicates when households adopt the app. Third, as suggested by the model-free evidence, household differences may exist that affect their stamp saving or redemption behavior. Though our control households are picked randomly and should not be different from the treated households, we use household-fixed effects to rule out any biases.

While the above considerations all pertain to measuring a common treatment effect of push messaging, we are additionally interested in *heterogeneous treatment effects.* As indicated in the conceptual part and Fig. 1, we seek to determine the *moderating* impact of two sets of observable consumer behavioral characteristics. First, we allow the treatment effect to vary by different levels of household spending prior to the program. Second, we allow it to be heterogeneous with respect to different levels of observed past participation during the program, i.e., stamp balance (or the cumulative number of stamps collected minus redeemed), and distance from the redemption threshold (or how many stamps a consumer is away from a possible redemption). These two moderators relate to each other, as the distance to the reward threshold is equal to the balance minus the number of stamps required to get a reward. But, while distance can at most be equal to the reward threshold, balance can be much larger and, together, the two variables allow to flexibly capture the impact of available stamps below and above the threshold. The variation in the three moderator variables stems from the between-household difference in the level of spending (collection) and redemption.

To determine these treatment effects, we require exogenous variation in our messaging variables. First, as mentioned earlier, a random control group receives no messages after week 3 of the program. Second, there is variation in the number of messages sent over time, and several individual messages were withheld from a random subsample of the treated consumers in different weeks. Finally, a fraction of the messages failed to be sent, but such failure is unintentional and thought to be related to technical issues. In sum, our push messages vary across households and time without being selectively targeted.

5.2. Model specification

To determine the effect of push messaging on consumers' spending (stamp collection), we use the following specification:

$$\begin{aligned}\text{SPEND}_{i,t} \;=\; & \alpha_i + \gamma_t + \beta_0 + \beta_1 \text{DIST}_{i,t} + \beta_2 \text{BALANCE}_{i,t} + \beta_3 \text{PREVREDEMP}_{i,t} \\ & + \beta_4 \text{ADOPT}_{i,t} + \beta_5 \text{LOGMESSAGES}_{i,t} + \beta_6 \text{LOGMESSAGES}_{i,t-1} \\ & + \beta_7 \text{LOGMESSAGES}_{i,t-2} + \beta_8 \text{DIST}_{i,t} \times \text{PREVSPEND}_i \\ & + \beta_9 \text{BALANCE}_{i,t} \times \text{PREVSPEND}_i + \beta_{10} \text{PREVREDEMP}_{i,t} \times \text{PREVSPEND}_i \\ & + \beta_{11} \text{ADOPT}_{i,t} \times \text{PREVSPEND}_i + \beta_{12} \text{LOGMESSAGES}_{i,t} \times \text{PREVSPEND}_i \\ & + \beta_{13} \text{LOGMESSAGES}_{i,t-1} \times \text{PREVSPEND}_i + \beta_{14} \text{LOGMESSAGES}_{i,t-2} \times \text{PREVSPEND}_i \\ & + \beta_{15} \text{LOGMESSAGES}_{i,t} \times \text{DIST}_{i,t} + \beta_{16} \text{LOGMESSAGES}_{i,t} \times \text{BALANCE}_{i,t} + \epsilon_{i,t},\end{aligned} \tag{1}$$

where the dependent variable $\text{SPEND}_{i,t}$ is the total amount of money spent at the retailer by household i in week t. Eq. (1) comprises three sets of explanatory variables. First, we include regressors that capture the dynamics of LP savings behavior (discussed in the conceptual part). The variables previous stamp balance ($\text{BALANCE}_{i,t}$), and distance to the reward-threshold ($\text{DIST}_{i,t}$) accommodate the points-pressure and goal-gradient mechanisms, respectively; while previous redemption ($\text{PREVREDEMP}_{i,t}$) captures possible rewarded-behavior effects. (We note that the correlation between balance and distance is not overly high ($\rho = -.45$), such that both variables can be included). We allow these saving dynamics to differ between more- or less-heavy buyers at the chain ($\text{BALANCE}_{i,t} \times \text{PREVSPEND}_i$, $\text{DIST}_{i,t} \times \text{PREVSPEND}_i$, and $\text{PREVREDEMP}_{i,t} \times \text{PREVSPEND}_i$). Second, as indicated above, to cleanly assess the push-message effects we also need to rule out several confounds. To control for a possible app effect, we include a dummy variable for app-adoption timing ($\text{ADOPT}_{i,t}$), and also allow its impact to vary with a household's prior spending at the chain ($\text{ADOPT}_{i,t} \times \text{PREVSPEND}_i$). Other household differences unrelated to push messaging (e.g., the fact that larger households may need larger baskets and thus 'collect' more stamps) are controlled for through household fixed effects (α_i). To control for 'message-unrelated' changes in spending in the course of the program (e.g. consumers collecting more stamps at the start) or other temporal confounds, we also use time fixed effects (γ_t).

The third set of regressors relate to the push message effects – our key variables of interest. The variable $\text{LOGMESSAGES}_{i,t}$captures the immediate effect of the number of messages received by the household in a given week, where we use a log transform to accommodate decreasing returns of multiple messages in the same week. We also include lags of

Table 4
Variable descriptions.

Variable	Description
Spending/Redemption variables	
$SPEND_{i,t}$	*Total spend (in 1000 Rupiah).* Measured as the total amount of money spent at the retailer by household *i* in week *t*. Equal to zero if no spending takes place that week.
$REDEEM_{i,t}$	*Stamps redeemed.* Measured as the number of stamps that are redeemed by household *i*in week *t*. Equal to zero if no redemption takes place that week.
$PREVREDEMP_{i,t}$	*Previous redemption.* Dummy variable (0/1) that indicates whether household *i* has made a redemption in the previous week.
Message variables	
$LOGMESSAGES_{i,t}$	*Ln_messages.* This variable is calculated as ln $(MESSAGES_{i,t} \times 100 + 1)$, where $MESSAGES_{i,t}$ equals the number of messages received by household *i* in week *t*.
$LOGMESSAGES_{i,t-x}$	*Lagged messages.* This variable is calculated as ln $(MESSAGES_{i,t-x} \times 100 + 1)$, for x=1, 2, where $MESSAGES_{i,t-x}$ equals the number of messages received by household *i* in the previous week $(t-1)$ or the week before that $(t-2)$.
Moderators	
$BALANCE_{i,t}$	*Previous stamp balance.* Measured as the stamp balance of household *i*, prior to (before the start of) week *t*. It is calculated as $BALANCE_{i,t-1} + COLLECT_{i,t-1} - REDEEM_{i,t-1}$. Stamp collection during a given week is reflected in the balance of week $t+1$.
$DIST_{i,t}$	*Distance to the reward threshold.* The number of stamps household *i* requires to reach the redemption threshold, prior to (before the start of) week *t*. In this case, we consider the target/threshold to be ten stamps, as this is the modal number of stamps required, for which they can redeem a reward. (See Web Appendix 5 for a robustness check). If at least ten stamps have been collected, or in other words, if the previous stamp balance is equal to at least ten, then distance to the target is equal to zero.
$PREVSPEND_i$	*Average spending prior to the program (in 1000 Rupiah).* Equal to the average (grand mean-centered) spending of household *i* in the weeks prior to the start of the program.
Controls	
$ADOPT_{i,t}$	*Adoption timing.* Step dummy that is equal to 1 from the app adoption time of household *i* onwards, 0 otherwise.
γ_t	*Week fixed effects.*
α_i	*Household fixed effects.*
$CF_{i,t}$	*Correction factor.* Included in the 'number of stamps redeemed' layer of the hurdle model, to capture the correlation with the redemption incidence layer (see McFadden & Dubin 1984); it is measured as $\left(1-\widehat{P}_{i,t}\right) * \frac{\ln\left(1-\widehat{P}_{i,t}\right)}{\widehat{P}_{i,t}} + \ln\left(\widehat{P}_{i,t}\right)$, where $\widehat{P}_{i,t}$ is the predicted redemption incidence probability for household *i* in week *t*.

this variable to capture the post-message effect due to purchase acceleration or carryover. In addition to these main treatment effects, we bring in several moderators. To determine whether messaging is more or less effective depending on consumers' stamp balance or distance, we include their interactions with the same-period message variable ($BALANCE_{i,t} \times LOGMESSAGES_{i,t}$, $DIST_{i,t} \times LOGMESSAGES_{i,t}$, where we mean-center distance and balance). To assess whether message response differs between more vs. less heavy buyers at the chain, we allow the immediate and lagged message variables to interact with the household's (mean-centered) average spending before the program ($PREVSPEND_i$). For an exact specification of each of these variables, see Table 4.

As mentioned before, in-store spending directly results in stamp collection, for app users. Moreover, because our interest is in the impact of push messaging within the program (and not in the program effect as such), we estimate the spending model on observations during the program only – excluding pre-program weeks. This way, we get a cleaner assessment of household fixed effects within the program and, thus, of the additional influence of push messaging.

To measure the effect of push messaging on stamps-redemption behavior, we use a hurdle specification with mostly the same regressors as in Eq. (1), in which the first layer captures the probability of redemption incidence for household *i* in week *t* ($REDEMPINC_{i,t}$) through a binary-logit specification, and the second layer captures the number of stamps redeemed given incidence ($REDEEM_{i,t}$).[11,12] This hurdle model naturally accommodates the (many) weeks in which a household does not redeem and allows us to predict the 'previous redemption incidence' regressor for our dynamic simulations (see also below). We estimate the two layers sequentially but, following McFadden and Dubin (1984), capture their interdependence through a correction factor in the second layer (see Table 4 for details).

We use the sandwich estimator of variance (White-Huber standard errors) to accommodate heteroscedasticity.

6. Results

To build up the results, we consider four nested specifications (which all include fixed effects): (1) a basic model with the main effects of distance, balance, and previous redemption, but no impact of push messaging, (2) a model with current and

[11] We note that the main effect of $PREVSPEND_i$ is subsumed in the household fixed effects. We include household fixed effects in the spending model, and in the 'number of stamps' layer of the hurdle model but not in the redemption incidence (logit) layer (where we include the PREVSPEND variable instead), because this would create a dynamic panel-regression problem.

[12] Because the number of stamps for different rewards are not multiples of one another, and consumers can (thus) redeem almost any number of stamps on a given occasion, a count model is not feasible here, and we can use a continuous (linear) specification for the number of stamps redeemed given redemption incidence.

Table 5
Estimation results.

	Spending	Redemption Incidence	# of Stamps Redeemed
$BALANCE_{i,t}$	−2.432***	0.0137***	0.319***
	(0.201)	(0.000441)	(0.0292)
$DIST_{i,t}$	−2.531***	−0.185***	−0.686*
	(0.433)	(0.00364)	(0.350)
$PREVREDEMP_{i,t}$	30.74***	1.256***	4.777*
	(1.783)	(0.0108)	(2.163)
$PREVSPEND_i$		0.00150***	
		(0.0000423)	
$ADOPT_{i,t}$	111.2***		
	(3.123)		
$LOGMESSAGES_{i,t}$	2.679***	0.125***	0.515*
	(0.437)	(0.00356)	(0.241)
$LOGMESSAGES_{i,t-1}$	−1.800***	0.0391***	0.0241
	(0.257)	(0.00340)	(0.0945)
$LOGMESSAGES_{i,t-2}$	−0.192	0.0106***	0.0731
	(0.252)	(0.00292)	(0.0485)
$BALANCE_{i,t} \times PREVSPEND_i$	−0.00227***	−0.0000152***	−0.0000734*
	(0.000326)	(0.000000506)	(0.0000358)
$DIST_{i,t} \times PREVSPEND_i$	−0.00449***	0.000137***	0.000259
	(0.00128)	(0.00000472)	(0.000284)
$PREVREDEMP_{i,t} \times PREVSPEND_i$	0.0446***	−0.000333***	0.0000332
	(0.00844)	(0.0000323)	(0.00104)
$ADOPT_{i,t} \times PREVSPEND_i$	0.269***		
	(0.0255)		
$LOGMESSAGES_{i,t} \times PREVSPEND_i$	0.0228***	0.00000740	0.0000505
	(0.00186)	(0.00000811)	(0.000156)
$LOGMESSAGES_{i,t-1} \times PREVSPEND_i$	−0.00803***	−0.0000200*	0.000132
	(0.00157)	(0.00000839)	(0.000147)
$LOGMESSAGES_{i,t-2} \times PREVSPEND_i$	−0.0155***	−0.0000360***	−0.000248*
	(0.00157)	(0.00000716)	(0.000122)
$LOGMESSAGES_{i,t} \times BALANCE_{i,t}$	−0.0685*	−0.000468***	0.00335
	(0.0304)	(0.0000871)	(0.00281)
$LOGMESSAGES_{i,t} \times DIST_{i,t}$	−0.578***	0.00497***	0.0272
	(0.0778)	(0.000744)	(0.0171)
$CF_{i,t}$			−2.467
			(2.020)
Constant	144.4***	−3.146***	5.038
	(4.988)	(0.0526)	(7.819)
Observations	651056	636981	93211
R^2 (pseudo-R^2)	0.042	(0.169)	0.168

All models include time fixed effects, the (linear) models for spending and # stamps redeemed also include household fixed effects, the (logit) model for redemption incidence does not. Standard errors in parentheses. Reported R-squares net of fixed effects.
* $p < 0.05$.
*** $p < 0.001$.

lagged push messages added, (3) a model that also adds households' pre-program spending, and (4) the 'full' model with the message-distance and message-balance interactions added to the previous specification. Web Appendix 2 gives the results for all models. Below, we discuss the estimates of the full model, reported in Table 5.

6.1. Spending

Column 1 of Table 5 lists the results for store expenditure/stamp collection. Considering the control variables, we find that program-app adoption comes with an increase in spending and stamp collection (β = 111.2, p < .001) – as expected. Our results also corroborate the presence of pre- and post-reward saving dynamics. First, we find that a smaller stamp balance increases spending (β = −2.432, p < .001). This is in line with the points-pressure mechanism: at levels of balance well below the savings threshold, the prospect of having to reach that (still remote) threshold within a given time frame leads consumers to make an extra effort, but this pressure wears off as the number of available stamps grows. Second, the negative coefficient of distance (β = −2.531, p < .001) supports the goal-gradient hypothesis that closeness to the redemption target motivates consumers to save more stamps (Kivetz et al., 2006). Third, previous redemption positively affects spending (β = 30.74, p < .001), in evidence of the rewarded-behavior mechanism (Drèze & Nunes, 2011).

Our focus, though, is on the impact of mobile push messages. Column 1 of Table 5 shows that push messages positively affect households' in-store spending and, thus, stamp collection. Considering the main effect, we obtain a positive immediate

effect of push messaging on expenditures ($\beta = 2.679$, $p < .001$). Turning to the lagged-message variables, the effect is negative for the one-week lag ($\beta = -1.800$, $p < .001$), but insignificant for the two-week lag ($p > .10$). This suggests that, similar to a post-promotion dip, consumers compensate for the spending lift at the time of receiving the message by buying less in the subsequent week (but not after).

The interaction effects with pre-program spending are all significant, suggesting there is considerable heterogeneity in the effects of messaging along different levels of prior spending at the chain. Our implications section provides further detail.

Next, we report on how saving dynamics interact with messaging. As Table 5 shows, the available stamp balance slightly reduces the effect of push messages on spending ($\beta = -.0685$, $p < .10$). As for distance, we expected that a smaller distance to the reward threshold would strengthen the immediate effect of messaging, and our results confirm such a negative interaction ($\beta = -.578$, $p < .001$). To show how these effects play out in combination, Fig. W3.1 in Web Appendix 3 plots the spending impact of push messages as a function of the consumer's available stamps.

Finally, we consider how the effectiveness of immediate and lagged messages depend on a household's pre-program spending with the retailer. The results show that the positive effect of immediate messages on spending increases with pre-program spending ($\beta = .0228$, $p < .001$). Immediate message effects are thus smaller for less-heavy buyers, who visit the store less regularly or progress towards the reward threshold more slowly and may be discouraged when they receive messages. As for lagged messages, the negative interactions with pre-program spending (1 lag: $\beta = -.00803$, 2 lags: $\beta = -.0155$, $p < .001$) implies that heavy buyers at the chain exhibit larger post-message spending dips. This is consistent with a 'ceiling' effect: because these customers already allocate (most of) their grocery budget to the chain, they have less room for extra outlay, and may partly compensate extra expenditures due to push by spending less subsequently.

6.2. Redemption incidence and number of stamps redeemed

Columns 2 and 3 of Table 5 report on the results for redemption incidence and stamps redeemed. Like before, the main effects of saving dynamics are as expected. Both redemption incidence, and the number of stamps redeemed given incidence, increase as the distance to the reward threshold goes down (incidence: $\beta = -.185$, #stamps: $\beta = -.686$, $p < .001$) and stamp balance goes up (incidence: $\beta = .0137$, #stamps: $\beta = .319$, $p < .001$). Furthermore, we obtain a positive effect of previous redemption on current redemption incidence ($\beta = 1.256$, $p < .001$) and redeemed stamps ($\beta = 4.777$, $p < .001$).

The immediate effect of push messaging is positive and significant, for both incidence ($\beta = .125$, $p < .001$) and number of stamps ($\beta = .515$, $p < .10$). So, as anticipated, current messages contribute to program awareness and remind consumers about redemption possibilities. Whereas previous-week messages affected in-store expenditure negatively, we find that they enhance the propensity to redeem (1 lag: $\beta = .0391$, 2 lags: $\beta = .0106$, $p < .001$), but do not affect the number of stamps traded in ($p > .10$).

As for the moderating effects, we find a positive interaction coefficient between messages and distance ($\beta = .00497$, $p < .001$) and a negative interaction coefficient between messages and balance ($\beta = -.000468$, $p < .001$). Though surprising at first, these coefficients must be interpreted as adjustments to the already built-in interactions in the logit model, which are governed by the main effects of distance and balance.[13] Together, these coefficients *do* imply that, as expected, the message impact on redemption incidence increases with higher balance and decreases with higher distance – be it less so than would be dictated by the built-in interactions. Fig. W3.2 in Web Appendix 3 plots the message effects on redemption as a function of available stamps.

7. Implications

7.1. Setup

As the model results show, the impact of push messages on consumers' shopping activity and program participation is statistically significant. We also find that it comes about in several ways, making it difficult to gauge the magnitude of the total (direct and indirect) effects of sending push messages. First, there is a direct effect of messaging on in-store spending and on redemption. Second, lagged messages influence current spending and redemption. Third, the impact of push messages depends on consumers' stamp balance and distance to the reward, quantities that endogenously evolve over time. Finally, more vs. less heavy buyers differ in their message response. Therefore, to assess the effects of push messaging over time for customers with different levels of (pre-program) spending with the retailer, we use our estimates as inputs for dynamic simulations.

Specifically, starting from the first week of the program, we predict consumers' spending and number of stamps redeemed in the course of the program for alternative push plans, and compare the predicted spending and redemption trajectories. We do so for different customer profiles – i.e. lighter vs. heavier buyers at the chain. For each profile, we simulate 1000 trajectories (each corresponding to a set of draws for redemption incidence) and average the outcomes across those

[13] The marginal impact of a message on the redemption probability is P*(1−P)*(B0+B1*Distance+B2*Balance), where B0, B1 and B2 are the main- and interaction- coefficients of message, and P is the 'Baseline' redemption probability. Because P increases in balance and decreases in distance, and because it is typically below .5, P*(1-P) will also increase (decrease) in balance (distance). Hence, as long as (B0+B1*Distance+B2*Balance) is positive, the message impact will be higher for higher balance and lower distance.

Table 6
Impact of adding extra messages.

Consumer Type	Baseline spending (in IDR, across program weeks*)	Change in		
		Spending (in IDR, across program weeks)	Number of Stamps Redeemed (across program weeks)	Stamp Balance (by the end of the program)
Panel A: Actual push plan vs. control group				
Average	1,927,595	270,216	18.12	−10.73
Low Quartile	1,717,076	219,642	15.42	−9.30
Median	1,822,081	244,480	16.82	−10.11
Upper Quartile	2,336,295	377,954	22.45	−12.44
Panel B: One additional push message in week 4 of the program vs. actual push plan				
Average	2,197,821	2,710	0.085	−0.021
Low Quartile	1,936,718	2,298	0.079	−0.024
Median	2,066,561	2,475	0.081	−0.024
Upper Quartile	2,714,249	3,891	0.130	−0.039
Panel C: One additional push message in week 7 of the program vs. actual push plan				
Average	1,527,929	2,958	0.119	−0.062
Low Quartile	1,349,536	2,214	0.084	−0.043
Median	1,438,291	2,322	0.092	−0.042
Upper Quartile	1,881,481	4,122	0.174	−0.088
Panel D: One additional push message in week 10 of the program vs. actual push plan				
Average	928,264	2,502	0.187	−0.150
Low Quartile	818,568	2,182	0.159	−0.128
Median	873,202	2,183	0.162	−0.131
Upper Quartile	1,144,139	2,619	0.189	−0.149

* Baseline spending equals the total amount spent from the week of the change onward if consumers were not to receive the extra messages. Hence, in Panel A, baseline spending is the total amount spent from week 4 onward (in which the actual and control plan start to diverge) until the end of the program, by consumers who would not receive any messages during those weeks. In Panel B (C, D), baseline spending is the amount spent from week 4 (7, 10) onward, by consumers who would not receive the extra message in week 4 (7, 10). Note: 1000 IDR is approximately equal to 0.07 USD. These numbers are considerable for Indonesian standards.

trajectories. Moreover, to account for uncertainty in the estimated parameters, we repeat these simulations for 1000 draws from the parameter sampling distributions, to assess the significance of the differences between alternative push plans. Web Appendix 4 gives more details on the simulation procedure.

7.2. To push or not to push?

We first assess the impact of the retailer's current push plan, i.e. with the 'actual' number of push messages sent out during the program as depicted in Fig. 2. We do this by contrasting the predicted spending and redemption levels of consumers 'treated' with that plan, with those of consumers in a 'control' scenario (who, as also shown in Fig. 2, receive no more messages after the third program week).

We start by considering the average effects, displayed in the top row of Panel A of Table 6. As the table shows, the push plan has a significant positive impact on spending. Starting from week 4 (which is the week in which the 'actual' and 'control' plan begin to diverge), treated consumers spend 270,216 IDR more (or, +24,565 IDR more per week on average) during the remainder of the program. So, the push plan produces a substantial (14.02%)[14] lift in consumer expenditures.

We also obtain a positive impact on the number of stamps redeemed. Consumers treated with the push plan redeem 18.12 more stamps (or, 1.65 more stamps per week) throughout the remaining 11 program weeks than they would in the control scenario. This comes down to a doubling of stamp redemption (+105.8%).[15] It follows that on average, messaging 'closes' the redemption gap: the additional number of stamps redeemed (18.12) by far exceeding the extra number of stamps collected (i.e. the 'stamps' equivalent of 270,216 IDR, or 6.76 stamps). As such, by the end of the program, the push plan reduces the average consumers' stamp balance by about one fourth (−10.73 stamps, a 23.4% drop).

These numbers, however, pertain to the 'average' customer. Because the message effects depend on households' spending with the chain prior to the program, and because the pre-program spending distribution across households is right skewed,[16] we also consider the impact of the push plan for other customer types (i.e. with pre-program spending levels at the lower-quartile, median, and upper-quartile of the pre-program spending distribution). The results are again summarized in Panel A of Table 6.

[14] Obtained as the spending lift among treated consumers during the treatment weeks (i.e. week 4 to 14), divided by the (simulated) baseline spending by control consumers during those same weeks, or: 270,216 IDR/ 1,927,595 IDR=.1402.

[15] Obtained as the redemption lift among treated consumers in weeks 4 through 14, divided by control consumers' (simulated) baseline redemption in those weeks, or: 18.12/17.12=1.058.

[16] The average pre-spending level is approximately equal to the 75th percentile of the spending distribution. Hence, our average household is considered a high spending household.

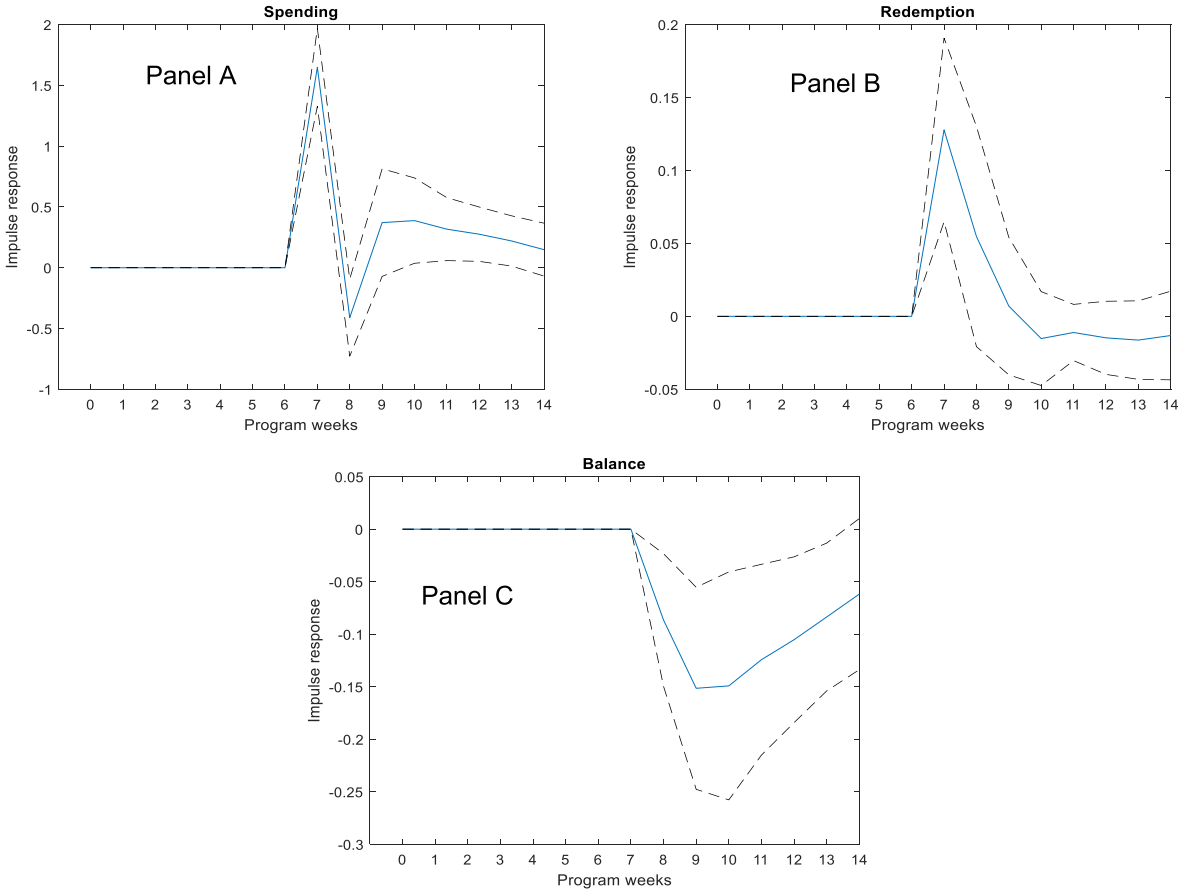

Fig. 4. Impact of an extra message in week 7 to the current push plan (average hh)

Zooming in on spending first, we note that less-heavy customers of the chain (the lower quartile of the distribution) already react positively to the push plan. Being sent more messages (from week 4 onwards), increases their expenditures at the chain by 219,642 IDR (a 12.8% lift) throughout the remaining program weeks. More 'heavy' buyers at the chain show larger expenditure increases in response to push messaging (+244,480 IDR and a 13.4% lift for median, +377,954 IDR and a 16.2% lift for upper-quartile customers). A similar pattern is observed for redemption: the current push plan leads upper-quartile customers to redeem more additional stamps in absolute terms (+22.45 stamps) than lower-quartile consumers (+15.42 stamps) and customers with median pre-program spending levels (+16.8 stamps).[17] The large difference between effects on spending versus redemption confirms that especially for customers who regularly visit the store and have saved many stamps, push messages can serve as a trigger to redeem.

Summarizing, we find that the current push plan has a strong positive impact on expenditures, and especially entices heavy customers to spend more at the chain. First and foremost, though, it stimulates program participants to trade in collected stamps for rewards – a phenomenon that is also particularly strong among heavy buyers at the chain.

7.3. How much push, to whom and when?

So far, we compared the actual push plan with the control scenario. However, managers can deviate from the current plan, and increase or reduce the number (and timing) of the messages as they see fit. To get a better feel for the impact of individual messages within the push plan, we consider the effect of adding one push message to the push plan actually used during the program. Because the message effects are dynamic, the return from an extra message may well depend on its timing. Hence, we report the results for three scenarios: one with an extra push in week 4 of the fourteen-week program, one with and extra message in week 7, and one with and extra message in week 10.

Fig. 4 shows the impulse response (i.e. the over-time impact of an additional message in week 7) for the average consumer, for spending (Panel A) and number of stamps redeemed (Panel B), across program weeks. The solid lines indicate the mean effects, and the dashed lines represent 95% confidence bounds (± two standard deviations around the mean). To

[17] All pairwise differences in spending and redemption lift between customer groups are significant at $p < .01$.

further clarify the dynamics, it also plots the evolution of the stamps balance – i.e. the number of stamps collected but not yet redeemed in a certain week (Panel C).

For spending, we see an initial spike in week 7 caused by the immediate effect of the extra push message in that week, a dip in the week after due to the negative lagged-message effects, and a return to positive values from week 9 onward. The plot for redemption also shows a spike at the time of the extra message (week 7), with a gradual return to the base level in subsequent weeks. The spending and redemption effects after week 9 are the result of the 'savings dynamics,' i.e., the impact of previous redemption, distance and balance. Indeed, the extra spending in week 9 follows from the 'rewarded behavior' created by the higher previous-week redemption, and from the ensuing reduction in stamp balance (see Fig. 4, Panel C), which provides a trigger to again step up collection. Yet, this same reduction in available stamps dampens the number of stamps redeemed in the course of subsequent weeks – only to gradually return to the base levels. As the plots show, positive redemption effects materialize especially in the short run, whereas increases in spending partly follow from lower balance in later periods.

Accumulating the impulse-responses in Fig. 4 across program weeks, we obtain the effects reported in the top row of Table 6, Panel B – the total change in spending and number of stamps redeemed from adding the extra message in week 7 for the average customer. The table also indicates the change in stamp balance by the end of the program, and breaks down each of these effects for lighter, medium, and heavier buyers at the chain. Table 6, Panel C provides the corresponding figures for a message added at a later time in the program (i.e. week 10 instead of week 7). Two additional insights emerge.

First, distinguishing lighter from heavier buyer segments, we observe a similar pattern as for the total push-plan effect: the effect of an extra message on spending and redemption is larger for consumers with higher levels of pre-program expenditures.

Second, the impact of an individual message depends on its timing within the program. Later messages, which reach consumers by the time they have built up a larger stamp balance, yield stronger redemption effects. In contrast, spending lifts are highest for mid-program messages. As can be seen from Fig. 4, this follows from three countervailing forces. First, push messaging increases spending within the same week (and entails a (smaller) spending dip in the week after). This produces a positive net effect especially early on in the program, when balance is still low. Second, push messaging also triggers redemption and thus reduces stamp balance, which stimulates customers to start collecting extra stamps (read: spending more) again. This redemption effect, and hence the associated upward spending shift, is bigger for messages sent later on. Third, however, for push messages near the end of the program, this positive 'uptake' effect is truncated in time. In all, this implies that from a spending perspective, messages around the middle of the program are most effective.

To rule out that this finding is idiosyncratic to the specific weeks in which we implemented the extra message, we consider alternative push plans with an additional message in still earlier or later weeks. The conclusion from these simulations comports with the findings above: adding more messages to the currently implemented plan enhances spending and redemption, especially for heavy buyers. However, the impact at the margin decreases for everyone. As for timing, the expenditure effects are highest for messages sent about halfway during the program. In contrast, later messages generate a stronger redemption lift.

7.4. Robustness checks

To validate the robustness of our results, we run several alternative models. (Web Appendix 5 gives detailed results). First, because households may collect a reward that requires fewer stamps than the 'modal' threshold (10), we re-estimate the model with the threshold set to five (the lowest reward value). This model yields a similar pattern of coefficients, but somewhat lower message effects (see Web Appendix 5, Table W5.1). Because there is only one reward with a 5-stamp requirement, and because this type of reward is not likely to be selected multiple times (unlike the other rewards, it is not a collectable), we retain the 10-stamps threshold in our main model.

Second, instead of using the log of messages, we consider an alternative specification, in which we allow for separate main-effect parameters for one, two and three messages per week (the range observed in our dataset). This model not only allows for decreasing returns but also for possibly negative effects as the number of messages goes up. We find that this model yields the same overall shape of effects as that with log messages: the impact of an extra message in the same week is still positive but becomes smaller and ultimately insignificant (see Web Appendix 5, Table W5.2). This confirms the presence of diminishing returns and suggests that for the message frequencies observed in our data, irritation is not an issue. Because the log-messages model is more parsimonious, we retain it as our main model.

Third, we check whether the dynamic message effects prevail if we include lagged spending as an extra explanatory variable. Though we find this variable to be collinear with previous redemption (leading to a coefficient shift for that variable), the dynamic message effects show the same pattern as before (see Web Appendix 5, Table W5.3, Panel B).[18]

7.5. Profitability

Push messages enhance consumer spending at the retailer. However, they have an even stronger effect on number of stamps redeemed. In many LPs – including the one studied here – such redemptions come at a cost to the retailer. Even

[18] We did not retain the model with lagged spending as our final specification, because the use of household fixed effects together with a lagged dependent variable might bias the individual coefficients of the spending model, and the use of Arellano-Bond estimators leads to very unstable outcomes (see also Web Appendix 5).

Table 7
Profitability of push messaging (actual push plan vs. control group)

Consumer Type	Baseline spending (in IDR, across program weeks*)	Change in		
		Margin on amount spent (in IDR, across program weeks)	Cost on redeemed stamps (in IDR, across program weeks)	Profit (in IDR, across program weeks)
Average	1,927,595	67,554.0	14,492.4	53,061.5
Low Quartile	1,717,076	54,910.4	12,337.1	42,573.3
Median	1,822,081	61,119.9	13,458.9	47,661.0
Upper Quartile	2,336,295	94,488.5	17,963.6	76,524.9

* Baseline spending operationalized as in Table 6. Note: 1000 IDR is approximately equal to 0.07 USD. Margin obtained as change in spending (see Table 6) times .25, a typical retailer margin for groceries. Cost obtained as change in redeemed stamps (see Table 6) divided by average number of stamps per reward (11.938) times average retailer loss per reward (9550.32 IDR) (Figures obtained from the company).

if consumers, in addition to handing in stamps, have to pay a certain amount to obtain a reward, this may not cover the wholesale price of the reward charged by the program operator to the retailer – implying that the retailer leaves money on the table for each redemption. The critical question thus remains whether the margin on extra sales from push messaging compensates for the additional loss on redeemed rewards, i.e. whether the push plan is actually profitable for the retailer.

We use average wholesale prices and consumer reward payments to the retailer[19] obtained directly from the program operator to explore this. Let STAMPS_REWARD be the average number of stamps redeemed per collected reward (in our case: 11.938 stamps), and LOSS_REWARD be the average 'net loss' for the retailer per reward, that is: the difference between the reward payments made by the consumer to the retailer, and the wholesale-price paid by the retailer to the program operator (i.e.: 9,550.32 IDR). Let ΔREDEEM (ΔSPEND) be the extra number of stamps redeemed (amount spent) due to push messaging (as displayed in Table 6), and MARGIN the retailer's average gross margin as a fraction of the sales price (which we set at .25, a typical retailer margin for groceries). The profitability of a push plan can then be approximated as:

$$\Delta\text{PROFIT} = \text{MARGIN} * \Delta\text{SPEND} - \text{LOSS\_REWARD} * (\Delta\text{REDEEM}/\text{STAMPS\_REWARD})$$

Table 7 presents the results of these calculations, for the actual push plan vs. the control. Though push messaging comes at a higher redemption cost, this is more than offset by the margin on additional consumer spending, in particular for heavy buyers. Counterfactuals reveal that this also holds for extra messages added to the actual plan, e.g. in weeks 4 and 10, or even more so in week 7. Hence, mobile push leads to incremental profit – especially among heavy customers, and for messages sent about halfway the program.

8. Conclusions, limitations and future research

Retailer LPs are pervasive in grocery retailing. However, even if they last only for a limited number of weeks, interest among program participants typically wears off, as evidenced by declining sales effects and failure to redeem stamps. Whereas traditional communication has not revealed very effective at maintaining program engagement, mobile push notifications sent to adopters of a loyalty-program app have been advanced as a potentially promising approach. Still, their effects are not clear-cut upfront. Though they may successfully act as reminders and maintain program salience, repeated messages to consumers' personalized mobile devices may lose effect or even become counterproductive. Moreover, if such messages primarily stimulate consumers to trade in already accumulated stamps rather than collect additional stamps, the net outcome for the retailer may not be profitable. Our study empirically documents these issues, using a very large-scale field experiment. In so doing, it contributes to the literature on LPs, but also on the, still young, body of work on the effects of mobile marketing.

8.1. Findings

Redemption. Mobile push messaging appears to be a very effective tool to stimulate redemption among program participants. This is not surprising, in view of previous evidence that LP participants often pile up stamps far beyond the redemption threshold. Push messages sent to consumers' personal mobile devices then remind them of 'earned' rewards, and provide a trigger to redeem those rewards. Our findings attest to this. Push messages produce a large immediate lift in the number of stamps redeemed, especially among heavy buyers (who can easily redeem stamps on their regular store visits), and more so when distance to the threshold is lower and stamp balance higher. This effect carries through to the next week because of rewarded behavior, to be followed by weeks in which consumers rebuild their stamp balance. The net result is a strong increase in number of stamps redeemed, especially from notifications sent later on in the program, and among the chain's large customers.

[19] These are "approximate" wholesale prices because not every redemption requires the same number of stamps or additional payment by the consumer. Though we do not have details on the exact reward(s) redeemed by individual households by redemption occasion, we know the relative occurrence of the different redemption types in the entire program. Our wholesale prices are weighted averages with these occurrences as weights.

Spending. Push messaging also substantially enhances customers' total stamp collection and spending at the chain. This effect materializes along several countervailing, routes. First, push messages have a positive immediate effect on spending – confirming that mobile messages can serve as a reminder of the program and of the gain from continued participation (Kivetz, 2003). This immediate effect is larger for customers with high pre-program spending – the group typically found to yield lower LP returns (Bijmolt et al., 2011) – who regularly visit the chain anyway and can thus easily 'act' upon the message. For lighter customers, this effect is smaller, possibly because these households find it harder to reach the reward thresholds. While we find evidence of decreasing returns (sending more messages in a given week leads to lower additional sales) we do not observe negative effects from extra push notifications, suggesting that among consumers who opted in for a program app these notifications do not easily trigger irritation or reactance – at least within the data range.

Second, we find a negative post-message effect, especially for the chain's largest customers. These customers may have less leeway to further increase their expenditures, and partly compensate for extra spending at the time of receiving the message by reducing their outlay subsequently. Third, beyond that period, the message effect turns positive again for some time. This follows from savings-dynamics mechanisms: rewarded behavior, and the lower stamp balance, give consumers an incentive to again step up their stamp collection efforts and make up for the spending dip. The outcome of these countervailing forces is a positive impact of push on spending, especially for messages sent halfway in the program. Hence, unlike (e)mailing programs for which modest results have been reported at best (Lewis, 2004; Dorotic, 2010), mobile push messaging reveals to be a powerful instrument to enhance stamps-saving behavior in LPs – and the expenditures that go with it.

Relative impact on spending and redemption. A notable observation is that push messaging has a stronger effect on redemption than spending, and thus closes the 'redemption gap.' Reminding consumers that they can earn a reward makes them more prone to redeem already collected stamps than to save extra stamps. For program operators, this is good news: it implies less money left on the table. For retailers, it increases the redemption cost associated with the program. Even so, our results reveal that this extra cost is more than compensated by the margin gain from additional spending, implying that push messaging is profitable for the retailer as well. For the program on hand, the net effect of the push plan amounts to 11% of the baseline margin on treated customers[20] – a non-negligible figure.

8.2. Management implications

Our results have important implications for managers involved in short-term LPs. They show that ongoing push notifications are a highly effective tool to ensure continued program engagement among participants enrolled in the program app, i.e., to make these consumers spend more and, especially, redeem more stamps. Given that push messaging significantly enhances the number of stamps redeemed, it can be used to tackle the phenomenon of un-redeemed stamps piling up – answering Dorotic et al. (2014)'s earlier call. Indeed, in the program under study we find that the current push plan reduces the number of unredeemed stamps by about one fourth – a very large effect. The redemption lift can substantially benefit the program operator and reward manufacturer – stakeholders that hardly received any attention in previous literature. For the retailer, the additional sales generated by the push program (in our setting: about a 14% lift) more than make up for the higher costs from increased redemption. Such a positive effect corroborates retailers' current tendency – as observed based on anecdotal evidence – to favor large numbers of messages.

Our results also provide retailers with guidelines on how to further improve the sales impact of their mobile push plans. Because the expenditure lift is particularly strong among heavy buyers (who typically account for a large portion of total turnover: in our setting: the top-25 customers are good for about 75% of retailer revenue), retailers have an interest in targeting those customers. As for timing: scheduling push halfway (rather than early or late in the program) may improve the sales-effectiveness of messages sent. Targeting and timing also shape the redemption effect of the program, but not in the same way. Messages sent in *later* program weeks (especially to heavy buyers), lead to the largest redemption increases. It follows that in setting up a push plan, the interests of retailers may not be aligned with those of the reward manufacturer or program operator – an issue especially relevant if the latter acts as the 'mobile-push captain'. Much will also depend on the program design in terms of consumer requirements to obtain rewards, and retailer payments to the program operator per redeemed reward. Typically, a balance is struck such that retailers can earn back the redemption costs through extra sales, while program operators also make a profit. Our results show that even if both parties benefit, push messaging shifts this balance in favor of the program operator and may become a relevant aspect in the total program design.

8.3. Limitations and future research.

While this study provides new insights, it also has limitations that entail opportunities for future research. First, our results pertain to one (temporary) program in one country. We considered the more typical setting of an 'investment' program, in which the retailer (at least partly) pays for the rewards. If the retailer earns money on redeemed rewards (as is the case for 'for profit' programs), the increased redemption caused by push messaging will translate into even higher gains. Also, while the processes that we uncover may generalize to most LPs, the magnitude of the specific effects may well be context-dependent. For instance, cultural differences may affect how many push messages are deemed acceptable by con-

[20] Approximated as the marginal gain from the push plan (53,062 IDR, Table 6), divided by gross profits (25% of baseline spending in Table 6 or 481,899 IDR).

sumers. Likewise, the exact nature of the reward or the redemption threshold may alter the timing of specific effects (e.g. the speed with which balance is restored). Last but not least, the absence of irritation effects may follow from the fact that in the program under study, consumers never received more than three push messages per week. Venkatesan and Kumar (2004) report that irritation effects may exist when communication frequencies are higher than those observed in our data.

Second, a broader question is whether our results generalize to permanent LPs where consumers collect points for expenditure that can be converted into rewards, e.g., frequent flier miles for free flights (Airlines), preferred customer programs for free nights or extra amenities (Hotels), customer cards for free coffee (Starbucks), etc.. While we expect the mechanisms to remain similar, the size of the message effects may change in a way that is not clear upfront. On the one hand, PLPs are bound to suffer more strongly from loss of salience and interest (Bijmolt et al., 2011), which could make for a stronger reminder effect of push messaging. On the other hand, push messages create a lower 'sense of urgency' in PLPs than TLPs. Even if consumers are reminded of the PLP and its rewards, there is less time pressure to redeem, and even less so to step up spending to qualify for redemption before saved points expire. Future research should document which of these forces prevail.

Third, our objective was to assess the impact of in-app push notifications on consumer engagement in a TLP. Hence, our population was restricted to card holders enrolled in the store's permanent LP (a prerequisite for TLP app adoption). Within that population, to ensure a clean comparison of test and control households, we estimated our effects on the early app adopters. Still, these constitute an important group, representing about 50% of all adopters, and spending 70% more per capita than late adopters.[21]

Fourth, we measured the impact of messages sent to consumers following their app adoption (and our test and control groups were random samples of app adopters), however, we do not know whether or when consumers read these messages, or possibly even disadopted the app (i.e. remove it from their mobile device). By the end of the program, we *do* observe lack of activity inside the app among certain households, but this may well be due to infrequent household purchasing rather than app disadoption. Indeed, such inactivity in final program weeks occurs mainly for households with low pre-program spending. Also, we do not find any link between such inactivity and the number of messages received by the household,[22] suggesting that app disadoption – if any – is not driven by push notifications.

Fifth, though we discussed possible causes of positive and negative message effects – including, for instance, reminder, discouragement and stockpiling effects – we could not directly quantify these intermediate process measures. Moreover, the push-message impact (over time) may differ depending on whether messages primarily remind consumers about available rewards, encourage them to save for those rewards, or stimulate them to redeem already collected stamps. For lack of information on the message content, we could not disentangle those effects, but leave that as a relevant topic for future study.

Finally, for lack of data, we only considered the impact of push messages during program weeks. To the extent that increased redemption also creates extra customer goodwill and commitment in the longer run – an aspect that we cannot measure here – our estimates of the push-message effects are conservative. We leave this as an area for future research.

Declaration of Competing Interest

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

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijresmar.2021.02.001.

References

Andrews, M., Goehring, J., Hui, S., Pancras, J., & Thornswood, L. (2016). Mobile promotions: A framework and research priorities. *Journal of Interactive Marketing, 34*, 15–24.

Andrews, M., Luo, X., Fang, Z., & Ghose, A. (2016). Mobile Ad Effectiveness: Hyper-Contextual Targeting with Crowdedness. *Marketing Science, 35*(2), 218–233.

Appel, G., Libai, B., Muller, E., & Shachar, R. (2020). On the monetization of mobile apps. *International Journal of Research in Marketing, 37*(1), 93–107.

Bart, Y., Stephen, A. T., & Sarvary, M. (2014). Which products are best suited to mobile advertising? A field study of mobile display advertising effects on consumer attitudes and intentions. *Journal of Marketing Research, 51*(June), 270–285.

Bellman, S., Potter, R. F., Treleaven-Hassard, S., Robinson, J. A., & Varan, D. (2011). The effectiveness of branded mobile phone apps. *Journal of Interactive Marketing, 25*(4), 191–200.

Bijmolt, T. H., Dorotic, M., & Verhoef, P. C. (2011). Loyalty programs: Generalizations on their adoption, effectiveness and design. *Foundations and Trends® in Marketing, 5*(4), 197–258.

[21] A caveat is that our spending data are recorded using consumers' loyalty card scans, while consumers can also purchase at the store without using their card. If push messages encourage card usage, we cannot disentangle that from spending increases and our spending impact may be somewhat overestimated. However, because much of our spending lift comes from an increase in quantity per visit, not visits, we do not expect this effect to be strong.

[22] We find a positive correlation between households' still being 'active' in the last program week and their pre-program spending (.233). Also, 'inactivity' is equally prevalent among households in the test and control group (21.29% vs. 19.43%; correlation between being active in week 44 and being in the control group: −.0102).

Bijmolt, T. H., & Verhoef, P. C. (2017). Loyalty programs: Current insights, research challenges, and emerging trends. In *Handbook of marketing decision models* (pp. 143–165). Cham: Springer.
Blattberg, R. C., Kim, B. D., & Neslin, S. A. (2008). Frequency reward programs. In *Database marketing* (pp. 549–578). New York, NY: Springer.
Bolton, R. N., Kannan, P. K., & Bramlett, M. D. (2000). Implications of loyalty program membership and service experiences for customer retention. *Journal of the Academy of Marketing Science, 28*, 95–108.
Bombaij, N. J. F., Gelper, S., & Dekimpe, M. G. (2020). How to design successful temporary loyalty programs. *Working Paper*.
Bombaij, N. J. F., & Dekimpe, M. G. (2020). When do loyalty programs work? The moderating role of design, retailer strategy, and country characteristics. *International Journal of Research in Marketing, 37*(1), 175–195.
Braun, M., & Moe, W. W. (2013). Online display advertising: Modeling the effects of multiple creatives and individual impression histories. *Marketing Science, 32*(5), 753–767.
Breugelmans, E., Bijmolt, T. H., Zhang, J., Basso, L. J., Dorotic, M., Kopalle, P., & Wünderlich, N. V. (2015). Advancing research on loyalty programs: A future research agenda. *Marketing Letters, 26*(2), 127–139.
Chauduri, M., Voorhees, C. M., & Beck, J. M. (2019). The effects of loyalty program introduction and design on short- and long-term sales and gross profits. *Journal of the Academy of Marketing Science, 47*, 640–658.
Danaher, P. J., Smith, M. S., Ranasinghe, K., & Danaher, T. S. (2015). Where, when, and how long: Factors that influence the redemption of mobile phone coupons. *Journal of Marketing Research, 52*(5), 710–772.
Danaher, P. J., Sajtos, L., & Danaher, T. S. (2020). Tactical use of rewards to enhance loyalty program effectiveness. *International Journal of Research in Marketing, 37*(3), 505–520.
Dorotic, M. (2010). *The influence of marketing instruments and rewarding on cardholders' behavior in coalition loyalty programs*. Groningen: University of Groningen, SOM Research School.
Dorotic, M., Verhoef, P. C., Fok, D., & Bijmolt, T. H. (2014). Reward redemption effects in a loyalty program when customers choose how much and when to redeem. *International Journal of Research in Marketing, 31*(4), 339–355.
Drèze, X., & Hoch, S. J. (1998). Exploiting the installed base using cross-merchandising and category destination programs. *International Journal of Research in Marketing, 15*(5), 459–471.
Drèze, X., & Nunes, J. C. (2011). Recurring goals and learning: The impact of successful reward attainment on purchase behavior. *Journal of Marketing Research, 48*(2), 268–281.
Dubé, J. P., Fang, Z., Fong, N., & Luo, X. (2017). Competitive price targeting with smartphone coupons. *Marketing Science, 36*(6), 944–975.
Fang, Z., Gu, B., Luo, X., & Xu, Y. (2015). Contemporaneous and delayed sales impact of location-based mobile promotions. *Information Systems Research, 26*(3), 552–564.
Fong, N., Fang, Z., & Luo, X. (2015). Geo-conquesting: Competitive locational targeting of mobile promotions. *Journal of Marketing Research, 52*(October), 726–735.
GfK (2015). Meer dan de helft van Nederland spaart mee voor huishoudelijke artikelen. Retrieved from https://www.gfk.com/nl/insights/news/meer-dan-de-helft-van-nederland-spaart-mee-voor-huishoudelijke-artikelen/..
Ghose Li & Liu (2019). Mobile targeting using customer trajectory patterns. *Management Science, 65*(1), 5027–5049.
Grewal, D., Bart, Y., Spann, M., & Zubcsek, P. P. (2016). Mobile advertising: A framework and research agenda. *Journal of Interactive Marketing, 34*, 3–14.
Gu, X., & Kannan, P. K. (2021). The dark side of mobile app adoption: Examining the impact on customers' multichannel purchase. *Journal of Marketing Research, 58*(2), 246–264.
Kim, B. D., Shi, M., & Srinivasan, K. (2001). Reward programs and tacit collusion. *Marketing Science, 20*(2), 99–120.
Kim, J. J., Steinhoff, L., & Palmatier, R. W. (2020). An emerging theory of loyalty program dynamics. *Journal of the Academy of Marketing Science, 48*, 1–25.
Kim, S. J., Wang, R. J. H., & Malthouse, E. C. (2015). The effects of adopting and using a brand's mobile application on customers' subsequent purchase behavior. *Journal of Interactive Marketing, 31*, 28–41.
Kivetz, R. (2003). The effects of effort and intrinsic motivation on risky choice. *Marketing Science, 22*(4), 477–502.
Kivetz, R., Urminsky, O., & Zheng, Y. (2006). The goal-gradient hypothesis resurrected: Purchase acceleration, illusionary goal progress, and customer retention. *Journal of Marketing Research, 43*(1), 39–58.
Koo, M., & Fishbach, A. (2008). Dynamics of self-regulation: How (un) accomplished goal actions affect motivation. *Journal of Personality and Social Psychology, 94*(2), 183.
Kopalle, P. K., Sun, Y., Neslin, S. A., Sun, B., & Swaminathan, V. (2012). The joint sales impact of frequency reward and customer tier components of loyalty programs. *Marketing Science, 31*(2), 216–235.
Kwong, J. Y., Soman, D., & Ho, C. K. (2011). The role of computational ease on the decision to spend loyalty program points. *Journal of Consumer Psychology, 21*(2), 146–156.
Lal, R., & Bell, D. E. (2003). The impact of frequent shopper programs in grocery retailing. *Quantitative Marketing and Economics, 1*(2), 179–202.
Lamberton, C., & Stephen, A. T. (2016). A thematic exploration of digital, social media, and mobile marketing: Research evolution from 2000 to 2015 and an agenda for future inquiry. *Journal of Marketing, 80*(November), 146–172.
Leenheer, J., Van Heerde, H. J., Bijmolt, T. H., & Smidts, A. (2007). Do loyalty programs really enhance behavioral loyalty? An empirical analysis accounting for self-selecting members. *International Journal of Research in Marketing, 24*(1), 31–47.
Lewis, M. (2004). The influence of loyalty programs and short-term promotions on customer retention. *Journal of Marketing Research, 41*(3), 281–292.
Li, C., Luo, X., Zhang, C., & Wang, X. (2017). Sunny, rainy, and cloudy with a chance of mobile promotion effectiveness. *Marketing Science, 36*(5), 762–779.
Liu, Y. (2007). The long-term impact of loyalty programs on consumer purchase behavior and loyalty. *Journal of Marketing, 71*(4), 19–35.
Liu, H., Lobschat, L., Verhoef, P. C., & Zhao, H. (2019). App adoption: The effect on purchasing of customers who have used a mobile website previously. *Journal of Interactive Marketing, 47*, 16–23.
Liu, Y., & Yang, R. (2009). Competing loyalty programs: Impact of market saturation, market share, and category expandability. *Journal of Marketing, 73*(January), 93–108.
Luo, X., Andrews, M., Fang, Z., & Phang, C. W. (2014). Mobile targeting. *Management Science, 60*(7), 1738–1756.
Mägi, A. W. (2003). Share of wallet in retailing: The effects of customer satisfaction, loyalty cards and shopper characteristics. *Journal of Retailing, 79*, 97–106.
McFadden, D. L., & Dubin, J. A. (1984). An econometric analysis of residential electric appliance holdings and consumption. *Econometrica, 52*(2), 45–362.
Meyer-Waarden, L. (2007). The effects of loyalty programs on customer lifetime duration and share of wallet. *Journal of Retailing, 83*(2), 223–236.
Meyer-Waarden, L., & Benavent, C. (2009). Grocery retail loyalty program effects: Self-selection or purchase behavior change?. *Journal of the Academy of Marketing Science, 37*, 345–358.
Mills, P., & Zamudio, C. (2018). Scanning for discounts: Examining the redemption of competing mobile coupons. *Journal of the Academy of Marketing Science, 46*, 964–982.
Minnema, A., Bijmolt, T. H., & Non, M. C. (2017). The impact of instant reward programs and bonus premiums on consumer purchase behavior. *International Journal of Research in Marketing, 34*(1), 194–211.
Narang, U., & Shankar, V. (2019). Mobile app adoption and online and offline purchases and product returns. *Marketing Science, 38*(5), 756–772.
Neslin, S. A., & Van Heerde, H. (2008). Promotion dynamics. *Foundations and Trends in Marketing, 3*(4), 177–268.
Osinga, E. C., Zevenbergen, M., & van Zuijlen, M. W. G. (2019). Do mobile banner ads increase sales? Yes, in the offline channel. *International Journal of Research in Marketing, 36*(3), 439–453.
Park, C., Park, Y., & Schweidel, D. A. (2018). The effects of mobile promotions on customer purchase dynamics. *International Journal of Research in Marketing, 35*(3), 453–470.

Phelps, J. E., D'Souza, G., & Nowak, G. J. (2001). Antecedents and consequences of consumer privacy concerns: An empirical investigation. *Journal of Interactive Marketing, 15*(4), 2–17.
Reichhart, P., Pescher, C., & Spann, M. (2013). A comparison of the effectiveness of e-mail coupons and mobile text message coupons for digital products. *Electronic Markets, 23*(3), 217–225.
Reinartz, W., & Linzbach, P. (2018). Customer loyalty and reward programs in retail in the digital age. In: Handbook of Research on Retailing, Gielens, K. & Gijsbrechts, E. (Eds), Edward Elgar Publishing, Cheltenham, UK, p. 296.
Rossi, F. (2018). Lower price or higher reward? Measuring the effect of consumers' preferences on reward programs. *Management Science, 64*(9), 4451–4470.
Shankar, V., Kleijnen, M., Ramanathan, S., Rizley, R., Holland, S., & Morrissey, S. (2016). Mobile shopper marketing: Key issues, current insights, and future research avenues. *Journal of Interactive Marketing, 34*, 37–48.
Sharp, B., & Sharp, A. (1997). Loyalty programs and their impact on repeat-purchase loyalty patterns. *International Journal of Research in Marketing, 14*, 473–486.
Stauss, B., Schmidt, M., & Schoeler, A. (2005). Customer frustration in loyalty programs. *International Journal of Service Industry Management, 16*(3), 229–252.
Stourm, V., Bradlow, E. T., & Fader, P. S. (2015). Stockpiling points in linear loyalty programs. *Journal of Marketing Research, 52*(2), 253–267.
Taylor, G. A., & Neslin, S. A. (2005). The current and future sales impact of a retail frequency reward program. *Journal of Retailing, 81*(4), 293–305.
Tsang, M. M., Ho, S. C., & Liang, T. P. (2004). Consumer attitudes toward mobile advertising: An empirical study. *International Journal of Electronic Commerce, 8*(3), 65–78.
Unni, R., & Harmon, R. (2007). Perceived effectiveness of push vs. pull mobile location based advertising. *Journal of Interactive Advertising, 7*(2), 28–40.
Van Heerde, H. J., Dinner, I. M., & Neslin, S. C. (2019). Engaging the unengaged customer: The value of a retailer mobile app. *International Journal of Research in Marketing, 36*(3), 420–438.
Venkatesan, R., & Kumar, V. (2004). A customer lifetime value framework for customer selection and resource allocation strategy. *Journal of Marketing, 68*(4), 106–125.
Wang, Y., Lewis, M., Cryder, C., & Sprigg, J. (2016). Enduring effects of goal achievement and failure within customer loyalty programs: A large-scale field experiment. *Marketing Science, 35*(4), 565–575.
Wang, R. J. H., Krishnamurthi, L., & Malthouse, E. C. (2018). When reward convenience meets a mobile app: Increasing customer participation in a coalition loyalty program. *Journal of the Association for Consumer Research, 3*(3), 314–329.
Zhang, J., & Breugelmans, E. (2012). The impact of an item-based loyalty program on consumer purchase behaviour. *Journal of Marketing Research, 49*(1), 50–65.
Zubcsek, P. P., Katona, Z., & Sarvary, M. (2017). Predicting mobile advertising response using consumer colocation networks. *Journal of Marketing, 81*(July), 109–126.