Available online at www.sciencedirect.com

ScienceDirect

Journal of Interactive Marketing 55 (2021) 118–145

JOURNAL OF INTERACTIVE MARKETING

www.elsevier.com/locate/intmar

# The Effectiveness of Triggered Email Marketing in Addressing Browse Abandonments

Marcel Goic [a,*] & Andrea Rojas [a] & Ignacio Saavedra [b]

[a] *Department of Industrial Engineering, University of Chile, Santiago, Chile*
[b] *School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania, United States*

**Abstract**

Triggered emails are personalized messages that are automatically sent as a response to specific actions or states of customers. Typical examples of this type of campaign include cross-selling recommendations, cart abandonment reminders, and re-engagement emails. Despite the widespread growth of these strategies, there has been no formal evaluation of their effectiveness. This paper investigates the impact of one type of triggered email campaign by using an experimental approach. We identify customers who had recently browsed the website of a multichannel retailer but had abandoned the process before making a purchase. Approximately half of the sample was randomly selected to receive automated emails with different configurations, while the other half receive no message at all. Comparison of the sales of these two groups indicates that browse abandonment emails have increase revenues in the online channel and in the triggered category. In terms of the design of the campaign, we found that the implementation of triggered emails plays an important role in their effectiveness. In this regard our result indicates that retargeting based on longer navigation histories is associated with larger conversions and that recommendations of wider assortments are associated with larger revenues.


*Keywords:* Direct marketing; Retargeting; Behavioral emails; Browse-abandonment; Triggered emails; Personalization

## Introduction

Electronic channels have been consistently developing in recent years, becoming a fundamental communication vehicle between retailers and their customers. These channels offer relevant content and have been integrated with traditional outlets to provide customers with a multichannel experience in which they can interact with a retailer not only in their stores but also through websites and mobile applications. In 2019, the number of email users reached almost 4 billion worldwide, and that number will continue growing.[1] Email is not only a massive form of communication, but it also enables firms to send personalized messages to their customers and generate timely evaluations of the messages' impact. There has been a recent tendency towards the personalization of content in retail, especially using electronic channels, such as websites and email (Ansari & Mela, 2003). *Event-based*, *behavioral messages*, or *triggered emails* correspond to personalized messages that are automatically sent as a response to specific actions or states of customers, and these messages add a new layer of personalization by defining specific events that help to identify the right time to communicate with customers. Some common examples of this type of campaign are reminders sent to customers when they abandon shopping carts or re-engagement messages when there has been a significant decrease in customer activity. Table 1 lists other examples of triggered emails.

There are good reasons to believe triggered emails can have large response rates compared to traditional emailing. First, the identification of the right timing to deliver a marketing communication to customers is an important driver of effectiveness (Li, Sun, & Montgomery, 2011). Second,

* Corresponding author at: Marcel Goic Beauchef 851, Santiago, Chile.
*E-mail addresses:* mgoic@uchile.cl (M. Goic), andrea.258@gmail.com (A. Rojas), ignacio.saavedra.m@gmail.com (I. Saavedra).

[1] Email Statistics Report, 2019–2023, the Radicati Group, Inc.

https://doi.org/10.1016/j.intmar.2021.02.002
1094-9968

Table 1
Examples of different types of automated contact with customers.

| Triggered email | Description |
|---|---|
| Confirmation | The objective is to welcome or thank a customer for a particular action performed on the website. |
| Order status | These act as a follow-up to customer purchases or processes initiated on the website, providing information on the state of their transactions. |
| Personal events | These are personal emails based on customer information, for example, sending birthday or anniversary greetings. |
| Cart and browse abandonment | The objective is to incentivize the customer to complete a transaction that was started in their previous session. |
| Cross-sell recommendation | These are emails with product recommendations based on the client's recent purchases. |
| Reorder | These are emails reminding the customers that they need to reorder a product. These are only valid for some product categories, which need to be bought regularly. |
| Reengagement | The objective is to encourage clients who have not visited the website or purchased recently to do so. |

triggered emails enable the identification of good prospects when most historical data is not very informative. For many product categories, customers' previous purchases are very good predictors of what they are going to buy in the future (Rossi, McCulloch, & Allenby, 1996). However, for some product categories, such as durable goods or other infrequently purchased items, the historical data at the customer level is insufficient to make a proper inference of purchase intention. For example, suppose that the firm wants to find customers to target a new price promotion for washing machines. By looking at transactional data, the firm will only find a limited number of customers with recent purchases in the category. Furthermore, customers who have purchased recently are probably the least likely to buy again in the category in the near future. On the other hand, customers who have recently been browsing the washing machine category on the firm's website are signaling they have a strong motivation to buy.

Given its potential advantages, many of the most active e-tailers have implemented automatic email responses to specific customer behaviors. According to the Email Marketing Industry Census 2018, 24% of marketers are using behavioral targeting based on web activity, and 39% plan to implement such measures in the short term. Additionally, 26% of e-marketers are using other types of multichannel triggers to communicate with their customers. In an exploratory study, we visited and subscribed to 16 retailers' websites and then recorded email responses to browsing products and adding items to the shopping carts. On average, the companies sent 16 emails per month, and approximately half of the sample sent some type of behavioral email (details of the responses are available in Appendix 1). These observations confirm that triggered emails are already in use by several retailers, and they indicate some common practices in terms of the timing and content of the messages. For example, we see that companies typically send triggered emails with short response times, waiting just a few hours between the customer's action (browsing or cart abandonment) and the execution of the email campaign.

Despite their potential effectiveness, triggered emails should be managed carefully. Consumers' perceptions of privacy impose limits on the amount of personal data that can be used to identify prospects (Nowak & Phelps, 1995) and how information is presented to retain consumer trust (Milne & Boza, 1999). Therefore, the timing and the type of information presented to customers in the messages sent can play an important part in the long-term success of a triggered email strategy. The need of registered e-mail addresses also limits the effectiveness of triggered campaigns. However, most retailers have email addresses for a large fraction of their customers, which combined with modern tools to identify internet cookies, results in a considerable number of customers that can be contacted using triggered emails. The company we collaborated with in this project knows the email address of a few million customers and they can identify between 30% and 40% of the visits to the website, which translates into a sizable number of customers that can be affected by a triggered email policy.

There is some anecdotal evidence suggesting that triggered emails have positive results compared to traditional campaigns. This evidence comes from e-commerce platform vendors and email service providers, such as Fresh Relevance, ExactTarget (now Salesforce), and Magento (a summary of the reported effectiveness is available in Appendix 2). The business cases present very favorable outcomes for triggered email marketing. In particular, for cart and browse abandonment emails, companies report consistent double-digit increments in open rates, the number of purchases, and sales revenues. Although these studies present valuable insights, the evaluations are not compared to concurrent controls; therefore, there is no assessment of causal relationships. In the context of behavioral targeting, this is a crucial concern because large response rates can be fundamentally driven by the selection of customers who would buy regardless of the firm intervention. Thus, the first contribution of this study is a formal evaluation of the effectiveness of triggered email marketing, isolating its effects from other confounding factors. A second contribution of this study is that it explores different implementations of the campaigns in terms of timing and content to characterize the heterogeneity of the effect of the treatment and understand under what conditions triggered emails perform better. Moreover, we evaluate the campaign effectiveness using a multichannel framework in which we consider conversions and sales in online and offline channels and decompose these metrics depending on own and cross-category effects.

In summary, despite allegedly good results in the industry, there is no formal evaluation of the impact of triggered email marketing in the literature. In this article, we propose an experimental evaluation of the effectiveness of this type of marketing activity for the case of a multichannel retailer, where we considering variations in some key design variables, such as timing, repetition, and the content of the value offering. Although the methodology we used can be applied to other event-based campaigns, in the empirical investigation, we restrict our attention to browsing abandonment events.

## Literature Review

We identified two streams of research that are relevant to the effectiveness of triggered email marketing campaigns: First, the use of emailing as a marketing communication tool and, second, the personalization of content in electronic environments.

### *Emailing in a Multichannel Context*

Even though several channels are involved in the implementation of triggered campaigns, we are especially interested in describing the previous work on customer responses to email marketing and the literature on browsing behavior as a mean of identifying customer purchase intentions. There is extensive work exploring how internet usage data can be used to predict behavior, ranging from the timing of service usage (Telang, Boatwright, & Mukhopadhyay, 2004) to the length of the browsing session (Bucklin & Sismeiro, 2003) and even to describing the shopping behavior across multiple websites (Park & Fader, 2004). In this investigation, we focus on specific browsing behaviors and assess the informative value of prospecting customers based on those behaviors.

Email was one of the first communication vehicles used extensively in the early stages of e-commerce, and, despite the emergence of several complementary communication channels, it continues to play an important role for many multichannel retailers. In fact, several industry reports indicate that email is not only one of the most widely used direct marketing channels, but it can also lead to increasing profitability (Zhang, Kumar, & Cosguner, 2017). While effective, on average, the previous literature has indicated that email responses are heterogeneous, depending on the customer profile (Wu, Li, & Liu, 2018).

The modern use of emailing is derived from previous types of direct marketing initiatives, such as traditional mailings and catalogs, on which extensive investigations have been conducted. For example, Bult and Wansbeek (1995) concluded that customers who have bought more in the past have higher response rates. Bonfrer and Drèze (2009) noted that email marketing presents some unique features that cannot be adapted directly from the traditional direct marketing literature. Additionally, emails also have distinctive characteristics in comparison to other digital channels. For instance, with respect to online display advertising, emailing provides more control for marketers in terms of the timing of the delivery and the context in which the message is displayed.

Another area of recent research is the management of email marketing programs, which is motivated by the increasing complexity of these systems. In this regard, Wu et al. (2018), propose a statistical methodology to quantify the effectiveness of an email marketing campaign, controlling for all relevant covariates. Kumar, Zhang, and Luo (2014) model the customer opt-in and opt-out rates of email programs and conclude that the marketing intensity indeed affects the opt-in and opt-out rates, but those responses are mediated by customer characteristics. More recently, Zhang et al. (2017) studied the customers' email opening and purchasing behaviors and find that customer activity in the email channel does not necessarily translate into larger sales.

Our analysis of triggered emails is performed in the context of multichannel retailing (Neslin et al., 2006). In essence, we are interested in evaluating the impact on sales in all the available channels when the customer browsing behavior is used to generate personalized value propositions. Thus, triggered email campaigns take advantage of one of the key benefits of adopting a multichannel perspective, namely, the complementarity of the customer data gathered through different channels to inform decisions about the marketing mix (Verhoef, McAlister, Malthouse, Krafft, & Ganesan, 2010).

### *Personalization and Automation*

Customer segmentation and targeting have been a focal point of investigations in direct marketing (Bult & Wansbeek, 1995; Levin & Zahavi, 2001). With the advent of digital channels and the development of new methods to learn about customer behavior, direct marketers can now automate the process of providing targeted content to each customer. Formally speaking, personalization is the decision regarding what marketing mix is suitable for each individual, based on previously collected customer data (Arora et al., 2008), and marketing automation is the process of specifying business rules and procedures to provide the marketing mix through computerized systems (Moriarty & Swartz, 1989). Although these concepts could be described separately, in the context of triggered emails, we only consider automated personalization. Taking into account that triggering events can occur at any time, for large customer databases, the process needs to be delegated to computerized routines.

By definition, triggered emails provide two fundamental components of personalization. On the one hand, they provide content that is relevant to each customer. In the case of browse abandonments, this includes the products they have personally chosen to visit on the website, products they have voluntarily abandoned in the shopping cart, or products that might be used together with other items they have recently bought. Previous research has shown that automatic adjustment of the marketing mix can add value by dynamically varying the products themselves (Golrezaei, Nazerzadeh, & Rusmevichientong, 2014) or the prices of those products (Zhang and Wedel, 2009). On the other hand, if properly calibrated, the content can be delivered at the right time, coordinating with the customer's evolution in the purchase decision process (Todri, Ghose, & Singh, 2020). Additional layers of personalization can also be added to these components, such as differentiating by age, gender, and intensity of use, among others.

A recent trend in online advertising is the use of retargeted advertising, where online ads are displayed to users after they visit the advertiser's website (Moriguchi, Xiong, & Luo, 2016; Sahni, Narayanan, & Kalyanam, 2019). Retargeted advertising is similar to triggered emails in that the targeting decisions are based on recent history of interactions, but there are a number of dimensions in which they differ. In comparison to triggered emails, retargeted ads are displayed on third-party platforms and, therefore, leave little room for elaborated communications. Additionally, retargeted ads have fewer restrictions in terms of the frequency at which they can be displayed, and, therefore, they tend to be shown very frequently and on many different websites. Finally, considering that retargeted ads require almost no customer information, they are less suited for relational marketing.

In general, the academic literature suggests that personalization provides several benefits to customers. It offers better communication and better preference matching to their needs, and it makes customers feel more important as individuals (Murray & Häubl, 2009; Vesanen, 2007). Personalization also brings benefits to marketers, as it generates higher response rates and profits, differentiation from other competitors, customer satisfaction, and customer loyalty (Postma & Brokke, 2002; Vesanen, 2007). However, personalization also brings relevant challenges. One of the major criticisms raised regarding personalization is that it might constitute an invasion of the consumer's privacy (Arora et al., 2008). Personalization necessarily implies showing customers that their transactional and demographic data is being used to generate content in a way that can be evaluated as invasive by some (Song, Kim, Kim, Lee, & Lee, 2016). Although some technical mechanisms are available to reduce privacy concerns, finding the proper balance between more detailed information, which leads to more effective recommendations, and the potential privacy concerns that come with this information is a major challenge for managers. In our context, for example, if a message is received right after a navigation event, the customer is reminded that the retailer is recording the list of pages they are visiting. Another challenge concerning personalization arises from its computational complexity (Montgomery & Smith, 2009). Although sophisticated statistical techniques can be used to fine-tune events triggering an automatic response, in our investigation, we use relatively simple rules that can be evaluated at no major computational cost. Unlike other personalization initiatives, such as shopbots (Smith, 2002) or adaptive websites (Urban, Hauser, Liberali, Braun, & Sultan, 2009) that require real-time responses, triggered emails do not require immediate responses and therefore they can be processed a few hours after the identification of the triggered event.

Regarding personalization in emails, Wattal, Telang, Mukhopadhyay, and Boatwright (2012) empirically evaluated consumer responses to different levels of personalization. They found that although most customers respond positively to product recommendations, some of them are negatively affected by the use of more personal information, such as greetings including their names. The use of email as a marketing channel has also motivated several recent academic investigations associated with personalization. For example, Song et al. (2016) study the effect of the personalization of email messages on customers' privacy risk perceptions. Similarly, Sahni, Wheeler, and Chintagunta (2018) also examine the effects of the personalization of email marketing, studying the impact on business metrics, such as sales and unsubscription rates. In a series of studies, they found that personalization leads to better performance. Moreover, they propose a number of mechanisms explaining why personalization might be associated with larger sales.

## Research Framework and Experimental Design

The main objectives of this research are to evaluate whether triggered emails can have a positive impact on sales and to identify the main drivers of their effectiveness. The evaluation of triggered email performance is challenging because it depends on whom the messages are targeted to and how the message is delivered. To organize the analysis, we propose a simple framework in which the effectiveness of a triggered email campaign depends on four broad groups of factors. Fig. 1 shows a schematic representation of these factors and lists some examples. In this figure, we highlight in bold letters the specific elements we consider in our empirical evaluation.

In this framework, we first consider the characteristics of the products that trigger the emails. We expect that behavioral emails can be more or less effective in boosting sales, depending on the length of the purchase cycles, the

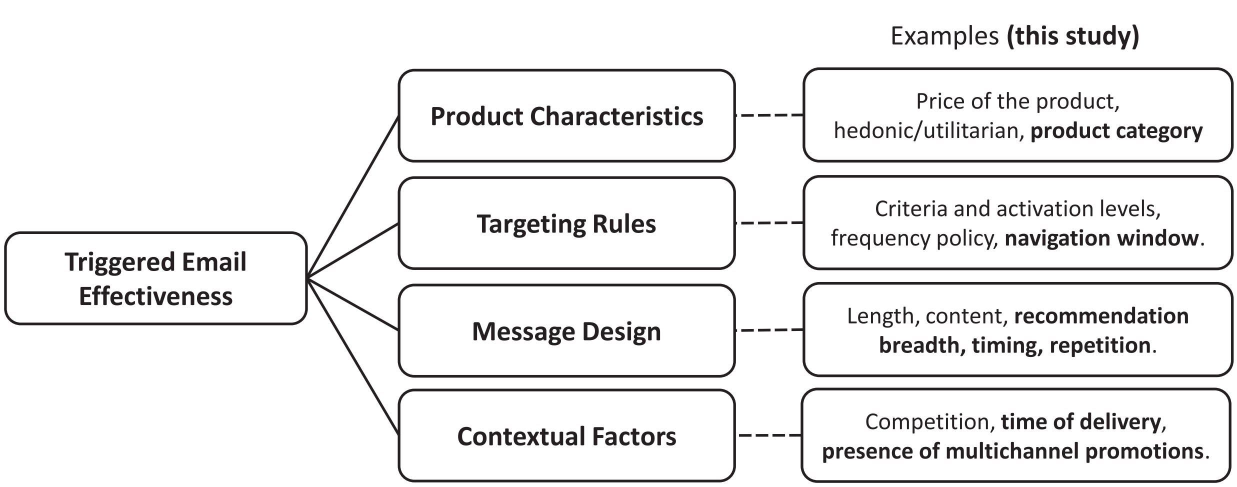


Fig. 1. Research framework for the evaluation of triggered email marketing.

involvement, and the competition. The response of an email can also depend on the specific characteristics of the product such as its price, obsolescence or even its hedonic or utilitarian nature (Chandon, Wansink, & Laurent, 2000). As the user can navigate in a wide variety of products, in this study we only control for the category the products belong to, which is readily available. In spite of only considering a small number of product categories (smartphones, LED TV, heaters, dryers, and washing machines), we consider important to control for them.

Next, we consider the rules used to determine which customers are eligible to receive the messages. Depending on the triggering event, there are several parameters to fine-tune the final list of prospects. For the case of browse abandonments, it is necessary to specify what metric is going to be used to measure the browsing activity (e.g., the number of sessions, time spent on the website, number of pages visited, or expected lift in conversions), and the minimum threshold required to trigger a message (at least four pages visited in the category, at least two pages more than the historical mean, etc.). In addition to these calibrations, some other filters could be applied to internalize some strategic business rules. For example, we might want to exclude customers who have already received a triggered message in the current week, those who have already purchased in the category in the past few days, or those that have a poor historical record of opening emails. The definition of the targeting rule is important because different behavioral patterns can provide stronger or weaker signals of customer interest in purchasing. In our study, we only consider messages sent to customers who concentrated at least 10% of their product views in the target category. However, we evaluate two levels for the realization of this type of triggering rule. In one scenario, we evaluate page views in the last two days, while in the other, we count the number of visits in the last four days.

The third group of factors is associated with the design of the message itself. There are many degrees of freedom to design more engaging communications with customers, but in our study, we analyze three design factors. First, we evaluate two criteria to select the products that are going to be recommended in the body of the email. In one scenario, we include products that are close substitutes for those products the customer visited the most. In the other scenario, we just recommend the most popular item in the category, regardless of the products that the customer has been visiting. If the events we used to trigger emails were associated with later stages of the purchase decision process, then consumers should be better served by a set of products that are more closely related to those they actively browsed on the website. On the other hand, general recommendations of the most popular products should work better for customers who are still learning about the variety of the product offering. Second, we consider the timing of the delivery of the messages. Previous research has shown that time is indeed a relevant component when deciding on personalized offerings (Zhang & Krishnamurthi, 2004). In our context, this is important because triggered emails are supposed to be synchronized with the evolution of the decision process of the customer. To generate sales, a marketing communication should be triggered as close as possible to the purchase decision. Waiting to send the message might result in the customer completing the process before the arrival of the email, when the information is no longer relevant. Although there exists evidence indicating that emails are mostly opened a few hours after being received (Bonfrer & Drèze, 2009), in the context of triggered emails, a delay between the identification of the event and the delivery of message might help to reduce the perception of an invasion of privacy (Blattberg, Kim, & Neslin, 2008; White et al., 2008). Thus, in our implementation of triggered emails, we impose a minimal temporal separation between the identification of a relevant navigation event and the delivery of the message to minimize the risk of the perception of invasiveness. More specifically, we consider two levels of delay, sending messages after two days for some users and after four days for others.

We also consider repetition as a relevant component of the design of the campaign. A common strategy used by direct marketers is retargeting customers who have received an email but have not opened it to increase the reach of the campaign. While most of the literature on advertising predicts that the performance of a campaign increases locally with the number of repetitions (Schumann, Petty, & Scott Clemons, 1990), there is also evidence suggesting that promotions wear out (Neslin, Powell, & Schneider Stone, 1995). To evaluate if repetition of triggered emails helps in increasing sales, in our evaluation we consider scenarios where we send a second email if the first is not opened.

A final group of covariates is associated with contextual factors. In general, triggering events can be identified anytime, and there might be some external conditions mediating the outcome of this promotional tool. For example, at the time of the event, the retail company might have some price discounts in the product category or the event might be triggered on a rainy day, thus affecting the daily sales (Steinker, Hoberg, & Thonemann, 2017). In our evaluation, we consider multiple scenarios that were executed in four weeks with relevant variations in a number of exogenous factors. Following previous findings in the literature, in our analysis, we explicitly consider if the products recommended in the email are on promotion (Reibstein, 2002; Vafainia, Breugelmans, & Bijmolt, 2019) and whether there was a banner for the corresponding category on the homepage of the retailer (Goic, Álvarez, & Montoya, 2018). Additionally, we consider dummy variables that indicate if the delivery occurred in the morning and if the national soccer team was playing in the World Cup, which might reduce the likelihood of shopping on that day.

In terms of the performance metrics, email campaigns are frequently evaluated using opening rates and click-through rates (Bonfrer & Drèze, 2009; Sahni et al., 2018). However, our identification strategy is based on an experimental approach, where a fraction of the potential recipients of the triggering email is not going to be exposed to them. Consequently, the opening and click-through rates are only available for the treatment conditions but not for the controls. Therefore, we evaluate triggered emails in terms of more direct metrics of performance, such as conversion rates and the revenue they generate. To have a more comprehensive understanding of the impact of triggered emails, in our analysis, we decompose sales based on the channel and category. Regarding the channel, we analyze whether the sales were completed in the online or offline channel. This is justified not only because the retailer we work with concentrated nearly 90% of its sales in brick-and-mortar stores, but also because previous research indicates that some customers might engage in *webrooming,* i.e.*,* a process in which product characteristics are researched in online channels but sales are materialized in offline stores (Gallino & Moreno, 2014; Verhoef, Neslin, & Vroomen, 2007). Regarding product categories, we measure whether a potential boost in sales is confined to the focal category that triggered the email or if there are some cross-category spillover into the sales of unrelated products.[2] Our evaluations are all made considering sales in the following three days after the delivery of the emails. This time frame is the default used by the web analytic platform the company uses for the evaluation of email campaigns.

## *Experimental Design*

The basic premise of triggered emails is that they can better select customers with a greater propensity to respond and that the information in such events can be used to personalize the message to increase its relevance to the customer. Despite the considerable potential of event-based marketing, communicating with customers who have already shown a significant interest in the product category might make it more difficult to generate any marginal change in their purchase behavior. Therefore, finding empirical support for a causal effect of triggered emails cannot be taken for granted and requires the definition of a proper baseline for comparison. To evaluate the differential impact of behavioral emails, we use an experimental approach, and we compare the response rates from customers for whom the event has been identified and a message has been sent against customers for whom the same type of event is also identified but there is no additional communication. In our experiment, we randomize the decision of who receives the email and therefore this comparison provides a clean evaluation of the impact of the treatment.

To evaluate if triggered emails are positively correlated with sales, we collaborated with a large multichannel department store in Chile. The company has a leading position with a market share close to 25% (50% in the online market) in the relevant markets, and, at the time of the experiment, approximately 10% of its sales were carried out through the online channel. In our study, we focus on browsing abandonment triggers in the appliances and electronic goods departments (LED TVs, Smartphones, washers, dryers, and heaters). The products in these categories receive a relatively large share of page views and the transactional data have relatively low explanatory power to describe short-term buying behavior.

The implementation of triggered email marketing campaigns brings several methodological challenges. Given that we were dealing with the active customer base of the company, there were some restrictions on the conditions that the firm was willing to explore. In terms of the product offering, they were willing to try different assortments but not different pricing levels.[3] Similarly, there were some strict rules regarding which customers could be contacted. For instance, customers who received any other promotional campaign in the previous week were not eligible to participate in the experiment. Second, the nature of events dramatically restricts the set of potential recipients of the campaigns. For example, to trigger an email inviting customers to complete a purchase in the washer dryer category, we need to observe customers with a minimum level of browsing activity in that category over the past few days. Consequently, on any given day, only a few hundred customers qualify to receive a triggered email in each product category. Therefore, we ran a series of experimental scenarios in several product categories for several days to accumulate enough data to draw meaningful statistical conclusions.

Based on the conceptual framework, in our experimental design we include a series of sequential scenarios with exogenous variation about how triggered emails are set. This design, where we simultaneously vary the treatment and the experimental context in which they are deployed, is instrumental to characterize variations of the magnitude of the treatment effect (Gerber & Green, 2012). The first variable we consider in the definition of the scenarios is the set of the products recommended in the email. In one condition, the product recommendations correspond to the most popular items according to recent web visits in the category. In the other condition, the recommendations are made among products sharing common attributes with the item the customer navigated to the most. Next, we consider the timing of the email that corresponds to the delay between the identification of the triggering event and the delivery of the message. Due to certain manual processing, during the study it was not possible to send emails until two days after the client visited products on the website. This constraint could affect the effectiveness of impulse purchases, but we hypothesize that it should be less severe for products in the electronics department. In our empirical evaluation, we tested sending emails two or four days after we detect that the customer exhibited an intense navigation pattern in a given product category. Finally, the scenarios vary according to the repetition strategy. For some scenarios, the customers receive only one message, regardless of their response, while for others, the customers who do not open the first message receive a second email with the same information as the original one.

To complete our definition of a triggered email scenario, we need to define which customers will be eligible to receive a triggered email. Unlike other experimental settings where all customers are eligible to be treated, in our case we are investigating how firms can use emails to communicate only with customers for whom their recent browsing activity

[2] To compute the sales in other categories, we only consider soft goods, such as clothing, bedding, and footwear, but we excluded more expensive product categories, such as furniture or electronics.

[3] Even though the firm has the ability to change prices relatively quickly, it considered that it would be difficult to justify to customers why some of them were receiving different prices.

indicates they are interested in a specific product category. In this regard, firms can decide more or less strict criteria to determine eligible customers. For instance, firms could send emails to every customer visiting at least one page in a given product category, or they could only send emails to the top 5% of customers ordered by the frequency of visitation in the last week. In our experiment, we used a fixed threshold of activity, but we vary if the evaluation is made only considering the browsing activity of the last two days or a longer navigation window of four days. With a shorter window, we might target customers with fresher browsing information, but the signal might be weaker.

By selecting different combinations of targeting rules, timing, content, and repetition of the message, we defined fourteen scenarios of browsing-abandonment triggered emails, as listed in Table 2. Levels for all factors are balanced providing enough variation for having no confound between main effects and first order interactions. Following common practice in experimental design with multiple conditions, we randomize the order in which we run these scenarios (Gunst & Mason, 2009).

Having defined the list of scenarios and a randomized order in which they are executed, for a given scenario we proceed as follows. We start by determining the list of eligible customers according to the targeting criterion in effect. Then, among the set of eligible customers we randomly select about half the customers to be assigned to the treatment and the other half to the control group. For the customers in the treatment group, we send emails with the content and delay in effect for that scenario. If the scenario considers it, we send a second email a few days later. Finally, to evaluate the impact of the intervention, we register the purchases of all eligible customers in the three days following the delivery of the first email. To illustrate the timing of the events, Fig. 2 displays the timeline of the execution of a given scenario. In our design, both the navigation window and the delay of the delivery are defined with respect to the moment where the eligibility of customers is determined. For example, for a scenario with a short navigation window, we evaluate if a consumer is suitable to receive a

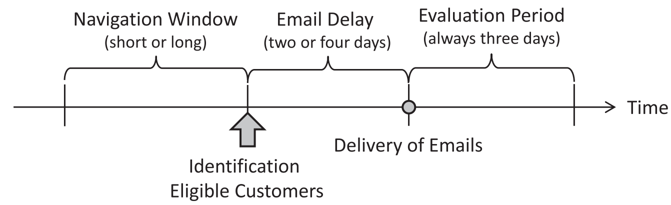


Fig. 2. Timing of the events for a given scenario.

triggered email by looking at browsing histories in the two days before the moment of the evaluation. Similarly, for a scenario with a short delay emails are sent two days after the identification of eligible customers.

The delivery of emails was conducted in a four-week period. The identification of eligible customers were typically conducten early in the week such that emails were sent between Wednesday and Friday. This is because we expect that some conversions were going to occur on the weekends, which is when the physical store of the retailer has its peak sales. For every scenario, we kept separate track of the product category where the customer was actively browsing and the randomization occurs within each product category. Table 3 lists the number of triggering events we identified for each scenario in each product category. For example, at the time of the execution of scenario S1, we identified 1907 customers who actively visited product pages in the smartphone category. From those, we randomly selected 953 customers and sent them an email with a narrow list of products recommended in the smartphone category. The other 954 customers received no message and were used as controls.

To implement the different versions of the campaigns, we prepared an email template, as illustrated in Fig. 3. In this template, each element of the campaign can be automatically modified depending on the characteristics of the customer and the triggering event. The structure of the template considers a main product, which is the item that the customer visited most on the previous days, but also a list of three recommended products. As we discussed earlier, this list was decided dynamically and depends on both the recommendation criteria

Table 2
Email campaign configurations.

| Scenario | Order | Product recommendation | Delay | Navigation window | Repetition |
|---|---|---|---|---|---|
| S1 | 1 | Narrow | 2 days | Short | No |
| S2 | 5 | Broad | 2 days | Long | No |
| S3 | 12 | Narrow | 4 days | Long | Yes |
| S4 | 7 | Broad | 4 days | Short | No |
| S5 | 13 | Narrow | 2 days | Long | Yes |
| S6 | 6 | Broad | 2 days | Short | Yes |
| S7 | 10 | Narrow | 2 days | Long | Yes |
| S8 | 2 | Broad | 2 days | Long | No |
| S9 | 11 | Narrow | 2 days | Short | No |
| S10 | 4 | Broad | 4 days | Long | Yes |
| S11 | 8 | Narrow | 2 days | Long | Yes |
| S12 | 14 | Broad | 2 days | Long | Yes |
| S13 | 9 | Narrow | 4 days | Short | No |
| S14 | 3 | Broad | 2 days | Long | No |

Table 3
Number of Triggering Events identified per scenario and category.

| Scenario | Smartphone | LED TV | Heaters | Dryers | Washing |
|---|---|---|---|---|---|
| S1 | 1907 | 1,383 | 234 | 530 | 470 |
| S2 | 662 | 509 | 149 | 140 | 163 |
| S3 | 861 | 883 | 201 | 112 | 156 |
| S4 | 599 | 548 | 143 | 103 | 176 |
| S5 | 727 | 552 | 173 | 102 | 145 |
| S6 | 575 | 730 | 116 | 68 | 126 |
| S7 | 463 | 683 | 136 | 102 | 116 |
| S8 | 1,042 | 700 | 136 | 0 | 180 |
| S9 | 400 | 540 | 142 | 60 | 63 |
| S10 | 569 | 477 | 95 | 73 | 74 |
| S11 | 379 | 506 | 163 | 83 | 122 |
| S12 | 473 | 527 | 164 | 94 | 148 |
| S13 | 388 | 506 | 143 | 87 | 166 |
| S14 | 299 | 210 | 54 | 50 | 49 |
| Total | 9,344 | 8,754 | 2049 | 1,604 | 2,154 |

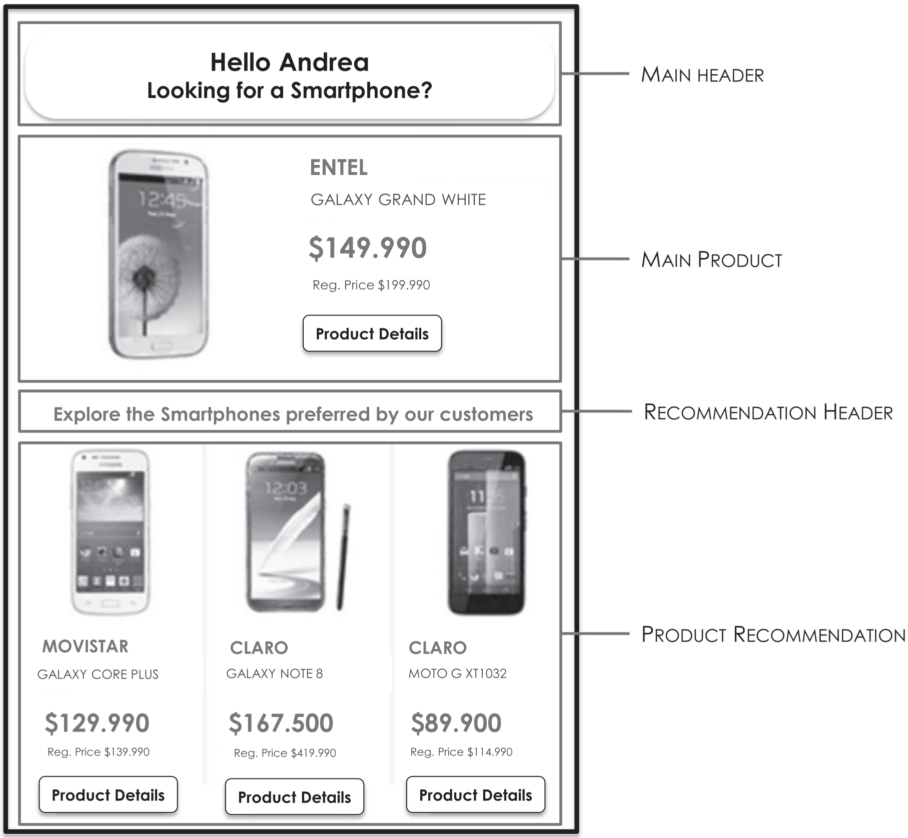


Fig. 3. Example of the template for the experimental triggered email.

(broad or narrow assortment) and the specific product the customer recently visited. The email subject for all emails is "[Customer name], find the [product] you are looking for," which is personalized with corresponding customer and product names.

Through the duration of the experiment, we identified 23,906 browsing abandonment events and sent 11,611 emails. However, in the statistical analysis, we discarded customers who, according to historical records, did not open any messages from the company in the last six months. When we include them in the analysis, the results are directionally the same but less significant. For each email sent, we observed whether the customer purchased or not from any department in any of the available channels and the revenue generated by that customer. In addition, we kept a record of a few complementary controls, such as the time of the delivery, whether the main product was on promotion, and if there were any internal banners in the corresponding product category.

## Modeling

### *The Effect of Triggers in Conversion and Revenues*

In our study, we observe two key performance metrics: Whether customers make a purchase in the evaluation period, and if they do, what is the revenue generate by their sale. These variables are closely related, and, therefore, we simultaneously estimate the effect of triggered emails on conversions and revenues. Econometrics literature refers to the data structure we observe as exhibiting *corner* or *zero-inflated* responses and offers alternative estimation approaches (Leung & Yu, 1996). In our case, we use a type-II Tobit model (Heckman, 1979), where we treat the continuous revenue variable conditional on a binary Probit model for whether or not the customer purchases from the retailer. For similar applications of this modeling approach see Fox, Montgomery, and Lodish (2004) and Srivastava and Kalro (2019). Formally speaking, let $y_i$ be a binary variable taking the value 1 if customer $i$ purchases from the retailer (0 otherwise) and let $r_i$ be the associate revenues. The regression model we used to estimate the effects of triggered emails is displayed in Eq. (1):

$$y_i = \begin{cases} 0 & \text{if } \alpha T_i + \theta_1' \mathbf{x}_{1i} + \varepsilon_{i1} < 0 \\ 1 & \text{otherwise} \end{cases}$$

$$\ln(r_i) = \begin{cases} 0 & \text{if } y_i = 0 \\ \beta T_i + \theta_2' \mathbf{x}_{2i} + \varepsilon_{i2} & \text{otherwise} \end{cases} \qquad (1)$$

In this model, the expression $\alpha T_i + \theta_1' \mathbf{x}_{1i} + \varepsilon_{i1}$ can be interpreted as a latent utility that results in a purchase if positive, and $\beta T_i + \theta_2' \mathbf{x}_{1i} + \varepsilon_{i2}$ is the observed log-revenue if that purchase takes place. In these expressions, $T_i$ takes the value 1 if customer $i$ received a triggered email (0 otherwise).

As $T_i$ is the experimental treatment, we denote it separately from every other covariate that we collapse in vectors $\mathbf{x}_{1i}$ and $\mathbf{x}_{2i}$. To complete the model, we assume the unobservables ($\varepsilon_{i1}$, $\varepsilon_{i2}$) are normally distributed (Amemiya, 1984) as described by Eq. (2).

$$\begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & \sigma \end{pmatrix}\right) \tag{2}$$

In our study, we consider different models depending on the set of covariates we include in $\mathbf{x}_{1i}$ and $\mathbf{x}_{2i}$. A common factor we included in all alternative specifications is category fixed effects in both equations. As we exclude the intercept, the category coefficients capture the baseline for the corresponding product category. The inclusion of these fixed effects is justified because purchase incidence and ticket value differ depending on product characteristics. Variations of the model result by adding customer characteristics and contextual variables that might moderate the effect of the triggered emails, and we include some of these alternative specifications in the appendices. Among customer characteristics we have age, a historical score associated with open rates for emails, and an internal segmentation from the company that classifies customers in *Registered*, *Online Buyers*, *Best Online*, and *Others*. Regarding contextual variables, we have two marketing mix variables that were not experimentally manipulated. First, we consider a dummy variable to indicate if the main product has a price discount. The second marketing variable is a dummy to denote the existence of banners or house ads associated with the target category on the homepage of the retailer (Goic et al., 2018). Finally, we included a binary variable indicating if the email was sent in the morning or the afternoon and another to denote if the soccer national team was playing in the World Cup on any of the days when we evaluate the impact of the treatment. Considering this is a popular event in the analyzed market, it may divert customers' attention, thus reducing the effectiveness of any marketing communication.

In theory, there is no restriction to the set of variables we include in each equation. Nevertheless, for practical purposes, it is useful to exclude some variables from the output equation as it alleviates the identification of the model from the nonlinearity of the inverse Mills ratio (Puhani, 2000). In our study, we selected the *Morning* and *World-Cup* dummies as the exclusion restrictions because we think they mostly affect the conversion $y_i$, but they are less likely to have a major effect on revenues $r_i$. Previous research has shown that the intensity of online shopping varies during the day (Goic & Olivares, 2019; Lee, Ha, Han, Rha, & Kwon, 2015) and therefore the time of the delivery of the emails might have an impact on the likelihood of converting. Similarly, if customers are occupied preparing for an important sports event, they might devote less time to shopping. These conditions mostly affect the availably of time and have no direct consequences on budget constraints, and, therefore, we only use them in the conversion equation. To evaluate how our results depend on these assumptions, we estimated an extensive battery of alternative specifications with different exclusion restrictions and verify that the main results remain (for details see Appendix 5).

## *The Design of More Effective Trigger Campaigns*

In this research, we are also interested in understanding what configurations of triggered emails might lead to better performance. In our experimental design, we have considered a series of scenarios with different triggering rules and content. To study the effect of these configurations, we run another regression model with a series of variables to characterize the type of treatment received as indicated in Eq. (3).

$$y_i = \begin{cases} 0 & \text{if } \alpha_N N_i + \alpha_{SD} SD_i + \alpha_R R_i + \alpha_W W_i + \theta_1' \mathbf{x}_{1i} + \varepsilon_{i1} < 0 \\ 1 & \text{otherwise} \end{cases}$$

$$\ln(r_i) = \begin{cases} 0 & if\, y_i = 0 \\ \beta_N N_i + \beta_{SD} SD_i + \beta_R R_i + \beta_W W_i + \theta_2' \mathbf{x}_{2i} + \varepsilon_{i2} & \text{otherwise} \end{cases} \tag{3}$$

In this equation, $N_i$ indicates the criteria used to select the list of recommended products included in the message. When $N_i$ takes the value 0, the customer received a broad assortment of the most popular items in the target category. For example, if the customer's most viewed product was a specific model of TV, then the recommended products correspond to the three TVs most viewed by other customers in the same period. Alternatively, when $N_i$ takes the value 1, the customer received a narrower assortment of products selected as the closest substitutes to those she visited the most. Following the previous example, in this case, the recommended products would be the three most viewed TV models, but with the same screen size and price range.

The variable $SD_i$ captures the time at which the message was delivered with respect to the moment where we evaluate eligibility. Here we consider two levels. The base level where $SD_i$ takes the value zero corresponds to a delay of four days after the identification of a large browsing activity in the target category. When $SD_i$ takes the value 1, it denotes a shorter delay of only two days between the identification of a browsing event and the delivery of the message. Next, the variable $R_i$ denotes whether the triggered strategy uses repetition or not. While multiple repetition strategies are possible in our experiment, we only consider up to one repetition for customers who did not open the first message. For the timing of the repetition, we consider waiting for one or two days from the original message, but we found no difference between them and therefore, we make no further differentiation in this regard.

The last design variable we consider in our regression model is devoted to characterizing the criterion to determine who is eligible to receive a triggered email. Here the dummy variable $W_i$ captures the length of the navigation window used to evaluate the activation of the trigger. Recall that customers were eligible to receive a trigger if we found that they had been actively browsing in the near past. More specifically, we considered customers who concentrated at least 10% of their product views in the target product category. However, to determine if a prospect qualifies, we need to define a time frame to compute his/her relative share of browsing. In our evaluation, we consider navigations windows from 3 to 5 days, but operationalize the variable as a binary indicator that takes the value $W_i = 1$ if the evaluation is conducted in a shorter window of 3 days (0 otherwise). As longer navigation windows are associated with larger browsing histories, they should lead to stronger signals of interest in purchasing the product. However, as they are more demanding, they capture a smaller fraction of potential buyers. Although there are multiple criteria to decide who would receive a triggered email, in our design we only distinguish how many days of navigation data is considered to decide eligibility.

## Results

We structure our results in three parts. First, we report model-free evidence of the effects of triggered emails and we present results from the main effect in conversion and revenues. Second, we report our evaluation about how alternative implementations of triggered campaigns can have an important impact on the performance of the campaign. Finally, we report conditional and total marginal effects to indicate the net expected impact of triggered emails in firms' revenues.

### *Main Effects*

Table 4 reports descriptive statistics for the data collected in the experiment broken down by experimental condition. We verify that, with the exception of a few small differences, the experimental conditions are very similar in the observables. To further guarantee the treated and control groups are indeed comparable, we apply a propensity score matching (Dehejia & Wahba, 2002), but it generates no meaningful difference in the main result (for more detail, see Appendix 3).

From the descriptive statistics, we observe that LED-TVs and smartphones triggered the majority of the emails. In terms of customer characteristics, many customers have not been classified in the company's internal segmentation and their ages predominantly ranged from 26 to 45 years. The marketing variables indicate that in 77.3% of the messages the main product has some price discount and in 80.6% of the emails the corresponding category had banners displayed on the homepage of the retailer. To address if these variables might induce a problem of multicollinearity, we compute variance inflation factors, and we found no problematic amount of correlation according to the most widely used rules to detect multicollinearity (James, Witten, Hastie, & Tibshirani, 2013). Detailed calculations are available in Appendix 4. In addition, in Appendix 6 we compare customer characteristics between conditions defining scenarios and we found no obvious differences either.

The last two rows of Table 4 provide aggregated statistics about the effect of the intervention. These numbers provide preliminary evidence that triggered emails generate a positive net impact on revenues for those who purchase (p-val = 0.022) but do not increase conversions (p-val = 0.57). To further understand the aggregated effects of triggered emails, we decompose the net impact of triggered emails by channel and category. In terms of channels, we distinguish between online transactions in the website and offline sales registered in the brick-and-mortar stores. In terms of the product category, we analyze sales belonging to the product category that generated the triggered in the first place (own-category sales) and those associated with any other category in the store (cross-category sales). These comparisons are displayed in Table 5 and provide aggregate evidence that the positive effect of triggered email marketing is mostly confined to the online channel and the product categories that triggered the emails. Recall that we sent emails to customers who were actively browsing in a given product category. Therefore, it is not surprising that the effect of the intervention translates into larger sales precisely in that category. Furthermore, as the emails are a firm's response to customer activity in the online channel, it is also expected that the effect is mostly capitalized by the same channel.

These results provide model-free evidence that triggered email marketing can boost revenues. To have a more comprehensive understanding of how different factors interact to increase sales, we present the results of the Tobit models that jointly analyze conversions and revenues while controlling for

Table 4
Descriptive statistics by experimental condition.

| | Treatment (triggered email) | Control (no email) | Total |
|---|---|---|---|
| LED-TV | 40.2% | 42.8% | 41.5% |
| Smartphone | 33.5% | 32.6% | 33.1% |
| Washing | 8.8% | 8.2% | 8.5% |
| Heater | 7.4% | 6.5% | 7.0% |
| Dryer | 10.6% | 9.9% | 10.3% |
| No-Age | 32.9% | 29.7% | 31.4% |
| Age 0–25 | 7.7% | 7.6% | 7.6% |
| Age 26–45 | 43.6% | 45.8% | 44.6% |
| Age 46 or more | 15.8% | 16.9% | 16.4% |
| No Segment | 44.2% | 40.9% | 42.6% |
| Registered | 12.8% | 13.6% | 13.2% |
| Online Buyers | 7.0% | 7.6% | 7.3% |
| Best Online | 16.9% | 18.5% | 17.7% |
| Others | 19.1% | 19.4% | 19.3% |
| Open rate score | 52.2 | 53.2 | 52.7 |
| Discount | 77.7% | 77.0% | 77.3% |
| Banner | 79.7% | 81.6% | 80.6% |
| Morning | 63.9% | 68.7% | 66.3% |
| World Cup | 39.6% | 45.9% | 42.7% |
| Conversion | 8.0% | 8.3% | 8.1% |
| Conditional Revenue [USD] | 189.34 | 120.49 | 155.74 |

Table 5
Model free comparison of net impact of triggered email marketing by channel and category.

| | Trigger | Control | Difference | p-val |
|---|---|---|---|---|
| Online | 8.481 | 4.490 | 3.991 | **0.036** |
| Offline | 6.589 | 5.455 | 1.134 | 0.497 |
| Own-category | 13.336 | 8.573 | 4.763 | **0.054** |
| Cross-category | 1.735 | 1.373 | 0.362 | 0.475 |
| Total sales | 15.071 | 9.946 | 5.125 | **0.042** |

Table 6
Tobit regression coefficients, for the average treatment effects.

| | Total | | Channel decomposition | | | | Category decomposition | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Purchase | p-val | Online | p-val | Offline | p-val | Own | p-val | Cross | p-val |
| *Conversion equation* | | | | | | | | | | |
| Trigger | −0.012 | 0.724 | 0.045 | 0.414 | −0.028 | 0.415 | 0.043 | 0.436 | −0.024 | 0.487 |
| LED-TV | −1.535 | <0.001 | −2.329 | <0.001 | −1.612 | <0.001 | −2.283 | <0.001 | −1.624 | <0.001 |
| Smartphone | −1.646 | <0.001 | −2.335 | <0.001 | −1.752 | <0.001 | −2.209 | <0.001 | −1.786 | <0.001 |
| Washing Machine | −1.429 | <0.001 | −2.149 | <0.001 | −1.537 | <0.001 | −2.105 | <0.001 | −1.533 | <0.001 |
| Heater | −1.498 | <0.001 | −2.432 | <0.001 | −1.544 | <0.001 | −2.546 | <0.001 | −1.522 | <0.001 |
| Dryer | −1.506 | <0.001 | −2.384 | <0.001 | −1.557 | <0.001 | −2.424 | <0.001 | −1.540 | <0.001 |
| Morning | 0.285 | <0.001 | 0.251 | <0.001 | 0.275 | <0.001 | 0.131 | 0.020 | 0.297 | <0.001 |
| World Cup | 0.042 | 0.205 | 0.100 | 0.075 | 0.011 | 0.755 | 0.149 | 0.007 | −0.002 | 0.959 |
| *Revenue equation* | | | | | | | | | | |
| Trigger | 0.188 | 0.144 | **0.577** | **0.087** | 0.048 | 0.700 | **0.370** | **0.064** | 0.036 | 0.675 |
| LED-TV | −0.131 | 0.896 | −3.595 | 0.311 | −0.213 | 0.839 | 0.100 | 0.970 | 0.610 | 0.373 |
| Smartphone | −0.040 | 0.969 | −3.840 | 0.277 | −0.103 | 0.924 | −0.192 | 0.941 | 0.521 | 0.467 |
| Washing Machine | 0.241 | 0.806 | −2.710 | 0.427 | −0.131 | 0.898 | 0.419 | 0.870 | 0.760 | 0.257 |
| Heater | −0.322 | 0.755 | −4.300 | 0.257 | −0.204 | 0.846 | −0.765 | 0.800 | 0.934 | 0.172 |
| Dryer | −0.510 | 0.617 | −4.891 | 0.178 | −0.328 | 0.754 | −1.700 | 0.551 | 0.734 | 0.279 |
| *Joint parameters* | | | | | | | | | | |
| Inv. Mills Ratio | 1.594 | 0.004 | 3.137 | 0.027 | 1.379 | 0.011 | 2.328 | 0.031 | 0.746 | 0.037 |
| $\sigma$ | 2.411 | | 3.647 | | 2.090 | | 2.494 | | 1.375 | |
| $\rho$ | 0.661 | | 0.860 | | 0.660 | | 0.933 | | 0.542 | |

other observables. We estimate the model using the two-step approach proposed by Heckman (1979) that is numerically stable and widely used in marketing research (see for example, Meire, Hewett, Ballings, Kumarand and Van den Poel, 2019 or Feng and Fay, 2020). Alternative estimation methods only produce small variations as is summarized in Appendix 5. Parameter estimates of Model (1) are reported in Table 6. In these regressions, we observe that the parameter associated with the Inverse Mills Ratio (IMR) is significant in all cases justifying the joint estimation of conversion and revenue equations.

Overall, the effect of triggered emails revealed by the Tobit model are consistent with descriptive statistics reported in Table 5. In terms of conversions, the coefficients are directionally as expected with positive values for online conversions in the targeted category. However, none of the effects in conversion is significant. Parameter estimates of the revenue equation exhibit significantly positive effects for the focal category (β = 0.370, p-val = 0.064) and in the online channel (β = 0.577, p-val = 0.087). Our finding that the positive effect on revenues is confined to online sales can be explained for two reasons. First, the event triggering the email is defined in terms of online activity and, therefore, all customers receiving the email can access and are familiar with the online channel. Second, the emails have direct links to the online channel, implying a very small cost of online shopping. On the other side, the association to physical stores is only indirect. Finding that the effect of the intervention is only significant in the targeted category can be explained because, in our context, customers receiving triggered emails have already been identified as browsing in a given product category, and they are likely to be at the final stages of their purchase processes with limited space for considering products in other categories.

In this previous analysis, in addition to the exclusion restrictions, we have only controlled by category fixed effects. Alternative specifications where we control for additional demographic and contextual variables produce no relevant change in the experimental coefficients. The only exception is that the effect on online revenue that is less significant (details available in Appendix 9). It is worth discussing that in these extended regressions, the marketing variables are almost never significant. We believe that this is due to the following reasons. First, we only observe binary indicators for price discounts and banners, and the majority of cases present some degree of promotion. Second, we only observe marketing variables concerning the main product, but the message also includes other products. As we measure the effect of triggers irrespective of the purchased product, it is plausible the effects dilute. Lastly, the customer we observe already expressed an interest in the product and therefore a significant part of the effect of the marketing mix already played a role in motivating the customer to navigate the website. Consider for example the case of price. We certainly believe that lower prices make products more attractive, but we are already looking at customers who evaluated prices to decide to browse actively on the website.

To derive the results of Table 6, we only use category dummies to define the corresponding baselines, but we did not consider that triggered emails can more effective in some categories than others. We report results with interactions between the treatment and the category dummies in Appendix 7, where the majority of the interaction coefficients are non-significant. Nevertheless, we find that triggered email could be detrimental in offline conversions of heaters and dryers and that triggered emails are particularly effective in increasing revenues for LED TVs.

We close the discussion of the results from Eq. (1), by pointing out that the indicator of the time of the day at which the messages are delivered is significant and sizable for all specifications we tried. In general, morning emails are associated with larger conversions. This is consistent with previous literature that shows important hourly seasonality in the online retailing (Goic & Olivares, 2019; Lee et al., 2015) and suggests that manager should be careful in determining the timing of the delivery of the messages.

*Evaluation of Alternative Implementation of Triggered Campaigns*

We can now discuss the effect of the design of the campaign in the outcome of triggered emails as described by Eq. (3). In our experimental setting, we only randomize whether we sent a triggered email to eligible customers, but we simultaneously vary the treatments and the experimental context in which they are deployed to characterize the heterogeneity in the treatment effect (Gerber & Green, 2012). These variations allow us to understand what implementations of triggered emails induce more favorable responses. Results for these analyses are shown in Table 7, where we once again display parameter estimates for conversion and revenue equations with the same set of controls used in Table 6. As a robustness check, we run a series of alternative specifications where we include additional covariates, but they produce no material changes in the effect of the design of the campaigns (see details in Appendix 10).

Regarding conversions, we found several design conditions with statistically significant impacts. For example, at the aggregated level, shorter navigation windows are associated with fewer conversions ($\alpha_N = -0.294$, p-val $< 0.001$). This is intuitive because using shorter windows to activate the 10% threshold we considered to trigger the emails are associated with fewer visits and, consequently, to a weaker signal of the customer propensity to purchase. Notice that unlike any other implementation variable that deals with the design of the communication, once a customer is identified, this variable is associated with the targeting rules defining who is eligible. From a managerial point of view, we should interpret the navigation window as informing us about what are the targeting rules that are more likely to correctly identifies customers that can be influenced through e-mail communications.

Parameters of the revenue equation show that repetition of messages, shorter delays, and broader product recommendations can boost sales at the chain level. Among them, the effect of recommendation criteria has the largest magnitude ($\beta_N = -0.887$, p-val $< 0.001$) with a significant effect in the focal category and in the online channel. This finding indicates that customers we consider eligible, in spite of showing a keen interest in a product category, might have not made a decision yet about the specific product to purchase. Therefore, they find value in receiving information about a more diverse assortment (Hoch, Bradlow, & Wansink, 1999). Additionally, as the

Table 7
Tobit regression coefficients, for the campaign design effects with additional controls.

| | Total | | Channel decomposition | | | | Category decomposition | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Sales | p-val | Online | p-val | Offline | p-val | Own | p-val | Cross | p-val |
| *Conversion equation* | | | | | | | | | | |
| Trigger | −0.008 | 0.812 | 0.040 | 0.478 | −0.023 | 0.513 | 0.036 | 0.510 | −0.019 | 0.594 |
| Narrow | **0.201** | **<0.001** | 0.036 | 0.609 | **0.242** | **<0.001** | **−0.141** | **0.034** | **0.310** | **<0.001** |
| Repeat | 0.027 | 0.538 | **0.248** | **0.001** | −0.056 | 0.236 | **0.159** | **0.034** | −0.026 | 0.588 |
| ShortD | −0.033 | 0.405 | −0.040 | 0.558 | −0.014 | 0.737 | **0.127** | **0.064** | **−0.083** | **0.053** |
| ShortW | **−0.294** | **<0.001** | **−0.156** | **0.032** | **−0.321** | **<0.001** | −0.066 | 0.332 | **−0.357** | **<0.001** |
| LED | −1.529 | <0.001 | −2.408 | <0.001 | −1.587 | <0.001 | −2.358 | <0.001 | −1.609 | <0.001 |
| Smartphone | −1.612 | <0.001 | −2.384 | <0.001 | −1.703 | <0.001 | −2.276 | <0.001 | −1.742 | <0.001 |
| Washing | −1.423 | <0.001 | −2.226 | <0.001 | −1.513 | <0.001 | −2.184 | <0.001 | −1.517 | <0.001 |
| Heater | −1.502 | <0.001 | −2.506 | <0.001 | −1.531 | <0.001 | −2.590 | <0.001 | −1.525 | <0.001 |
| Dryer | −1.494 | <0.001 | −2.451 | <0.001 | −1.529 | <0.001 | −2.493 | <0.001 | −1.520 | <0.001 |
| Morning | 0.142 | <0.001 | 0.120 | 0.077 | 0.141 | 0.001 | 0.118 | 0.071 | 0.128 | 0.002 |
| WCup | 0.165 | <0.001 | 0.273 | 0.001 | 0.102 | 0.037 | 0.177 | 0.022 | 0.156 | 0.002 |
| *Revenue equation* | | | | | | | | | | |
| Trigger | 0.157 | 0.218 | 0.429 | 0.127 | 0.028 | 0.833 | **0.358** | **0.080** | 0.015 | 0.872 |
| Narrow | **−0.887** | **<0.001** | **−1.877** | **<0.001** | −0.322 | 0.183 | **−1.238** | **<0.001** | −0.078 | 0.652 |
| Repeat | **0.378** | **0.016** | 0.331 | 0.353 | 0.078 | 0.675 | 0.287 | 0.206 | 0.035 | 0.781 |
| ShortD | **0.344** | **0.024** | **0.766** | **0.032** | 0.152 | 0.333 | **0.575** | **0.030** | −0.035 | 0.763 |
| ShortW | 0.304 | 0.285 | **1.263** | **0.002** | −0.184 | 0.613 | 0.309 | 0.254 | −0.019 | 0.938 |
| LED | −0.619 | 0.661 | −0.897 | 0.773 | −1.704 | 0.359 | −0.930 | 0.750 | −0.567 | 0.627 |
| Smartphone | −0.481 | 0.741 | −1.133 | 0.714 | −1.607 | 0.406 | −1.131 | 0.690 | −0.686 | 0.577 |
| Washing | −0.091 | 0.947 | −0.178 | 0.953 | −1.462 | 0.421 | −0.540 | 0.846 | −0.309 | 0.786 |
| Heater | −0.602 | 0.671 | −1.274 | 0.694 | −1.535 | 0.403 | −1.520 | 0.636 | −0.166 | 0.884 |
| Dryer | −0.867 | 0.538 | −2.123 | 0.507 | −1.683 | 0.358 | −2.825 | 0.361 | −0.357 | 0.753 |
| Inv. Mills Ratio | 1.871 | 0.019 | 2.087 | 0.107 | 2.219 | 0.026 | 2.723 | 0.021 | 1.404 | 0.026 |
| $\sigma$ | 2.522 | | 2.708 | | 2.612 | | 2.792 | | 1.745 | |
| $\rho$ | 0.742 | | 0.771 | | 0.849 | | 0.975 | | 0.804 | |

alternative is given by a collection of products that other customers already found popular, a better valuation of that set can be explained by the *wisdom of the crowd effect*, where the collective action of the customers generates good recommendations (Hertwig, 2012). The higher revenues reported for earlier email delivery is consistent with previous research indicating that in the context of retargeting, the effectiveness of advertising decreases as time passes from the visit that triggers the retargeting ad (Sahni et al., 2019).

It is worth noting that the estimates reported in Table 7 correspond to the change in conversion and revenues with respect to the base configuration of each level (broad recommendations, no repetition, long delays, and long navigation windows). To evaluate if a given level is better than the non-treatment alternative, in Appendix 11 we run a sequence of models when we compare one factor at the time with respect to the non-treatment condition, revealing that in some cases the effects of triggers are only driven by one level. For example, when we look at the delay, we found that short delays generate a positive impact on revenues, but emails sent after longer delays do not lead to additional revenues.

From a managerial point of view, the main takeaway from these results is that, when implementing triggered email marketing strategies, the design matters. For example, when considering all configurations in Table 6, we find that triggered email marketing has a positive but non-significant effect in online conversion ($\alpha_T$ = 0.045, p-val = 0.414). Results from Table 7 now indicate that the total effect in conversions of a triggered email that include repetition (with all other design variables in their base levels) would be the sum of the corresponding coefficients ($\alpha_T + \alpha_R$ = 0.040 + 0.248 = 0.284) and it becomes significant (p-val < 0.001). Thus, by properly calibrating how to select customers, repetition strategies, the content of the communication and the timing of the delivery, online marketers can significantly improve the impact of a triggered email marketing initiative.

### Marginal Effects

To understand the profitability of triggered email marketing, we need to compute marginal effects. As is pointed out by Vance (2009), the marginal effect for the Tobit model must be calculated using a nonlinear function of the underlying model parameters to correct for selection. In our case, the triggered treatment affects the selection and output equations, and, therefore, we estimate the effect on the purchase probability, the effect on revenues conditional on purchase, and the unconditional effect on revenues, as reported in Table 8.

Table 8
Marginal effects.

| | Purchase probability | | Conditional revenue | | Unconditional revenue | |
|---|---|---|---|---|---|---|
| | Mean | s.d | Mean | s.d | Mean | s.d |
| Online | 0.002 | 0.001 | 14.866 | 7.276 | 0.791 | 0.369 |
| Offline | −0.004 | 0.001 | 0.388 | 0.080 | 0.008 | 0.002 |
| Own category | 0.002 | 0.001 | 36.786 | 13.938 | 0.860 | 0.336 |
| Cross category | −0.003 | 0.001 | 0.159 | 0.020 | 0.001 | 0.001 |
| Total | −0.002 | 0.000 | 1.474 | 0.370 | 0.102 | 0.021 |

The computation of the first two columns is direct. The marginal effect in the selection equation is equivalent to the marginal effect in a probit model (Anderson & Newell, 2003). For the case of the conditional revenue, we use the exponential of the linear estimates corrected by the effect of the selection (Saha, Capps Jr, & Byrne, 1997). Finally, to compute the unconditional effect on revenues, we follow McDonald and Moffitt (1980) to argue that change in the (log) revenues can be decomposed in the change in the expected value of the revenues among the positive revenues, weighted by the probability of being positive (given a treatment) and the change in the probability of being having positive revenues, weighted by the conditional expected revenue (given no treatment). For a detailed explanation about how to compute marginal effects in our context, see Appendix 8.

All these marginal calculations depend on $\mathbf{x}_{1i}$, and, therefore, we compute them for each individual. In Table 8 we report the mean and standard deviations across all individuals. The intuition of the marginal effects is straightforward. Consider for example the case of online revenues. Here, triggered emails have a very small impact on increasing purchase probability in 0.002 on average. However, conditional on a purchase, the intervention increases revenues by US$14.9. Considering only a small fraction of customer purchase, the unconditional effect per email send is US$0.791. These averages consider all configurations including those that are less effective as described in Table 7. Firms interested in implementing triggered emails are likely to adopt those configurations that exhibit more promising results. For example, according to our estimates, the counterfactual scenarios where all emails were sent using narrow recommendations would imply an unconditional lift in online revenue of US$1.12 per email and if the emails were sent with a short delay, those revenues would be as high as US$1.74 per email.

## Discussion

Triggered email marketing is an automated process that sends messages to customers as a response to specific actions taken by website visitors. Compared to traditional emails, triggered messages can be cost-effective, allowing for communication with the right customer at the right time. However, to evaluate their marginal value we need to consider that customers contacted with triggered emails have already shown a strong signal of their interest in the product category, and, therefore, they are likely to purchase anyway. Compared to an experimental control, we find that triggered emails are indeed associated with larger revenues. When translating the impact into monetary value, we find that each triggered email is associated with a mean incremental revenue of US$0.791. Furthermore, this figure can be even larger if the campaigns are optimized in terms of repetition, activation threshold, and product recommendations criteria. Notice, however, that the reach of this retargeting tool is limited to customers who already exhibited a noticeable interest in a given product category.

The positive influence of triggered emails is driven by sales occurring in the promoted category and in the online channel. This is expected because the emails have direct links to visit products in that category and in the online channel. In general, a positive effect in the online channel is consistent with the evaluation of other personalized marketing communications. For example, Sahni et al. (2019) found that retargeted ads can increase website visits by more than 14%. Moreover, Moriguchi et al. (2016) found that the effectiveness of this type of ad depends on the content (in both cases, they do not evaluate the cross-category nor cross-channel effects).

In terms of channel substitutions, previous research indicates that retail channels might be substitute or synergetic depending on the specific setting and customer motivations (Kollmann, Kuckertz, & Kayser, 2012). On the one hand, Breugelmans and Campo (2016) show that price promotions in the online channel reduce offline category sales in short term. On the other hand, Dinner, Van Heerde, and Neslin (2014) show that online advertising can boost offline sales. In our findings we did not find strong evidence for channel substitution. Although this could be partially explained by the need of larger sample size (Gordon, Zettelmeyer, Bhargava, & Chapsky, 2019), we believe this is explained because the emails we study in this research are only triggered after the customer show strong signals of being close to the end of their decision process. In this case, we consider less likely that marketing stimulus generate an increment of the total number of purchases.

In our experiment, we varied several attributes of the campaign to shed light on how triggered marketing should be implemented to be more successful in boosting sales. For instance, we considered different criteria for recommending products, the delay between the browsing event and the execution of the campaign, and the possibility of sending a second email to customers who did not open the first one. These results show that a mere reminder is not enough and that the design of the triggered emails can play an important role in their effectiveness. For example, here we find that product recommendations based on the most popular items lead to more revenues than those where we recommend close substitutes of the most visited products. We also find that repetition of emails is an additional source of revenue and that messages sent relatively soon after the identification of the trigger are more effective. As we only compared two against four days of delay, our experiment is unable to determine the performance of even shorter delays. While we expect they can be more effective, shorter delays should be managed carefully because they might create more privacy concerns on the consumer side. Certainly, our evaluation only considers some of the factors that can play a role in the effectiveness of triggered emailing. Firms interested in implementing this strategy would need to fine tune the configurations that better work for their particular cases.

The aforementioned positive impact of triggered emails should be taken with caution. First, in our analysis, we discarded customers with no recent activity in the email channel. When we include them in the analysis, the effects are directionally the same but not all of them are significant to the 95% level. This result indicates that channel affinity should be considered when implementing behavioral communications. In essence, triggered emails use information concerning customer behavior in one channel to motivate a direct communication in another. However, this channel compatibility is not guaranteed (Zhang et al., 2017). Despite a growing number of multichannel customers, a significant proportion of customers still concentrate their interactions with the retailers through a limited number of communication channels. An important challenge for leveraging information from multiple channels is the ability to identify customers who are responsive to this strategy. Second, we find little evidence of triggered emails affecting conversions. We believe this is at least partially explained because we concentrate our analysis in the electronic department, where it is relatively difficult to influence conversion decisions. To confirm this belief, it would be necessary to expand the study to a broader set of product categories, which we consider an interesting question for future research.

Our investigation has other limitations that should be considered when attempting to generalize its results. In the introduction, we noted that many different events can be used to trigger a marketing action. In this study, we only analyzed browse abandonment recovery emails. We posit that different events should not only imply different behaviors but should also be treated in a different manner. For example, a cross-selling recommendation might require different timing than the browse abandonment we analyzed here. The effectiveness of triggered emails might also depend on the retailer implementing them in terms of its competitive position. In our case, we collaborated with a retailer with a strong position in the relevant markets. For a market leader, event-based marketing might be less effective than for smaller retailers. If no triggered email is sent to customers actively searching in a product category, we think they are more likely to end up purchasing from the retailer with a dominant position. Thus, a prompt reaction from smaller retailers might be more critical for converting customers. The multichannel strategy of the retailer is also important. As behavioral marketing requires detailed tracking of customer activity, retailers with a crafted multichannel strategy can better identify relevant events. For example, if the customer is identified as actively browsing in a product category in more than one channel, he/she is giving a much stronger signal that he/she is getting close to making a purchase decision. Similarly, if the retailer offers an integrated multichannel experience, there are more opportunities to give the customer a good product offering at the right time and in the right place.

In this study, we have explored a limited set of variables that the retailers can manage to accommodate customers' needs. For example, we used a simple rule to identify email recipients, but some retailers could come up with more sophisticated rules to fine-tune the right targets. More specifically, the intensity of visits triggering an email could depend on the visit frequency of the customer, the likelihood of the customer opening an email, or the time that has elapsed since a previous purchase. More sophisticated implementations could even consider targeting rules bases on uplift modeling (Gubela, Lessmann, & Jaroszewicz, 2020). Similarly, the time at which the message is sent, or even the communication channel, can be decided at the customer level based on transactional data. In our experiment, we only attempted two alternative levels for the delays and found that a delay of two days is preferable to one of four. However, the optimal waiting time might depend on the product category. For example, for impulse buy items, for which we expect short decision processes, it might be optimal to trigger a communication before two days, while for high-involvement items, it might be better to wait longer than the four days we explored here. In our analysis, we also control for a limited number of product characteristics and adding more features can enhance our understanding of triggered emails. A notable example would be the price of the main product. While we do not keep records of these prices, it would be interesting to know how triggered email marketing works depending on price tiers. Furthermore, a more comprehensive list of controls would allow us to shed light on the variation of the number of customers who qualify to receive an email.

Another limitation is that we analyze triggered emails in isolation and not in comparison to alternative retargeting tools. A detailed evaluation of this sort can help managers to decide how to combine different strategies to achieve more effective marketing communications. A final limitation of our study is the evaluation horizon. In this study, we have only evaluated the campaigns in the short term, but triggered emails can have effects in the long term. On the negative side, the use of event-based marketing can increase privacy concerns and consumers might be more reluctant to make browsing information available to retailers, or, worse, not buy from them. On the positive side, the proliferation of event-based marketing might lead to a more rational system for communicating with customers, limiting the contact to only those instances in which they are interested in a specific value offering. In summary, the use of event-based marketing has the potential not only to drive sales but also to enable smarter communications with customers. To take full advantage of this paradigm, firms need to carefully design process workflows and business rules to decide the optimal conditions in which to contact customers. In this study, we used browse abandonment as the trigger and email as the communication vehicle. However, the administration of events has a larger scope. An effective event-based strategy should consider many customer events and multiple communication channels to provide a meaningful contextual experience.

## Acknowledgments

The first author gratefully acknowledge partial funding by Fondecyt (Project No 1191745) and the Instituto Milenio para la Investigación de Imperfecciones de Mercado y Políticas Públicas IS130002.

## Appendix A. Examples of Companies Implementing Triggered Emails

Table A1
Triggered email benchmark.

| Company | Cart abandonment emails | Browse recovery emails |
|---|---|---|
| Amazon | Yes, 1 email within 24 hours. | Yes. The email addresses the customer by name and shows the product viewed. |
| Dafiti | Yes, the first email is sent 1 hour later, showing the abandoned product, and a second email is sent 24 hours. Later, including a discount or free shipping. The second email is sent only if the first is not opened. | Yes. The first email is sent within a few hours, showing the browsed products and other recommendations. A second email is sent if the first is not opened, offering a 5% discount (in the subject line). Similarly, a third email is sent if the second is not opened, offering a 10% discount. The company also triggers emails when lowering prices on products in the customer's wish list. |
| Macy's | No | No |
| J. Crew | No | No |
| Crate and Barrel | No | Yes. The first email is sent the next day, displaying the abandoned product. A second email is sent 3 days later if the first email is not opened, once again displaying the product. A third email is sent if the second one is not opened, offering other product recommendations. |
| The Home Depot | Yes, 1 email 3 days later or 8 days later. | No |
| L.L. Bean | Yes, 1 email within 24 hours. | No |
| Booking | Yes (abandoned reservation), 1 email within 24 hours. | Yes. The email addresses the customer by name and is sent within 24 hours. The email shows the browsed destination and includes recommendations for similar destinations. If the customer does not purchase, a similar email is sent 48 hours later. |
| Ebay | Yes. The email is sent 7 days after the product is abandoned. | Yes, the email is sent 5 days after browsing the product. |
| Lululemon | No | No |
| Walmart | Yes. The email is sent 7 days later and shows recommendations for similar products. | Yes. The email is sent 6 days later. |
| Target | No | No |
| Best Buy | No | No |
| Uniqlo | No | No |

## Appendix B. Triggered Email Business Cases

Table A2
Triggered email business cases.

| Company | Email solution | Results |
|---|---|---|
| The Book Depot Partnership | Real-time website recommendations<br>Browse recovery emails<br>Abandonment recovery emails | 27% increase in sales |
| Julep (beauty) | Browse recovery emails<br>Improve email personalization<br>(Abandonment recovery emails already implemented) | 3.3% increase in sales |
| AlexandAlexa (kids department store) | Improvement in cart recovery emails (more personalized, from 24 hours. delay to<br>real-time response) | 69% increase in email orders<br>71% increase in email conversion<br>257% increase in email revenue |
| Spa Boutique | Improvement in cart recovery emails (from 1 contact to 3 contacts with delays<br>of 30 minutes., 24 hours. and 72 hours.) | 5.5% increase in recovery rate<br>8% increase in revenue |
| Wasserstrom (food service) | Browse recovery emails<br>Abandonment recovery emails | 5% increase in sales (in 90 days) |
| Cottages4you (holiday cottages) | Browse recovery emails<br>Abandonment recovery emails (1 hour. delay) | 957% email ROI |
| Moss Bros. (menswear) | Abandonment recovery emails with product recommendations (real-time) | 80% email open rate |
| 7dayshop (technology store) | Abandonment recovery emails (real-time)<br>Email personalization using shopping data | 20% email conversion<br>6% increase in revenue |

## Appendix C. Preprocessing Using Matching Methods

During the execution of our study, in each scenario and category, half of the eligible customers who satisfied the triggering rules were selected to receive an email with a personalized assortment of recommended products, while the other half of the customers received no message and were left as the control. This random assignment implies that the treatment and control groups should be well balanced in terms of all observables and, therefore, any difference in the purchase behavior can be attributed to the triggered emails. However, our final sample is collected in a sequence of scenarios conducted on different days, which might lead to minor differences associated with variations in the nonexperimental variables. To further guarantee that experimental and control groups are indeed comparable, we preprocessed the data using a propensity score matching approach (Dehejia & Wahba, 2002). To build the propensity scores, we use customer demographics (age, gender, and socio-economic group), internal customer segmentations (e.g., “Online Buyers”) and other contextual observables (for example, whether the message was sent in the morning or in the afternoon). The variations of the matching algorithm regarding expanding or contracting the set of covariates generate very minor differences. In terms of the matching algorithm, our results are based on a nearest-neighborhood approach. As a robustness check, we also performed genetic matching. This approach requires much more computational time, and in our particular case, it provided no further gain in balance. Overall, using matching to weight our sample leads to a small improvement in the balance of the dataset as we explain next.

### *C.1. Matching Quality Assessment*

To assess the effectiveness of the matching procedures used, we start by comparing the distributions of the propensity scores before and after matching. Fig. A3 displays these distributions. A number of observations are worth mentioning from these figures. First, the distributions of the treated and control groups are fairly similar before the matching, thus confirming the relative success of our randomization. Moreover, the distribution is concentrated around 50%, which is indeed consistent with our randomization strategy of selecting half of the sample to be treated. We notice, however, that pre-processing the experimental data with a matching approach leads to a small improvement in the balance of the distributions, especially in their upper tails. This observation is confirmed when moving beyond the propensity score distribution to analyze each individual covariate, as displayed in Table A3. When comparing the variables between the treatment and control groups, we observe that they are fairly balanced before treatment, with an average absolute difference of 2.1%. Nevertheless, the use of matching can help to further reduce the differences to an average mean difference of only 1.2%. Moreover, the differences are not only small on average, but every single variable has a small deviation, with a maximum of 3.2%. When using raw data or matched samples produce no material changes in the effects reported in the main document.

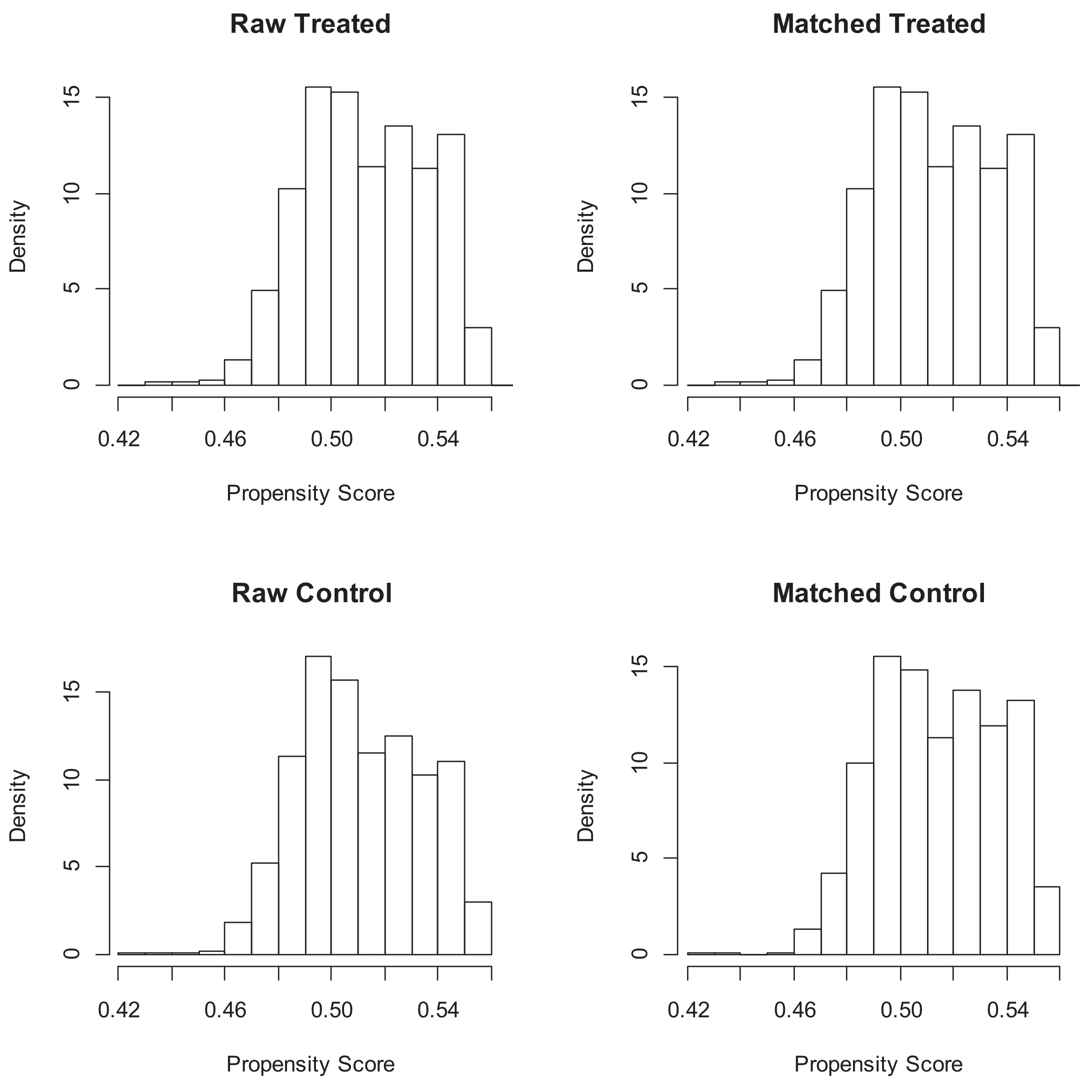


Fig. A3. Distribution of the propensity scores before and after matching.

Table A3
Avg. sample characteristics before and after matching.

| Variables | Control | Before matching | | After matching | |
|---|---|---|---|---|---|
| | | Trigger | \|Δ\| | Trigger | \|Δ\| |
| Age 0–25 | 7.6% | 8.1% | 0.5% | 7.7% | 0.1% |
| Age 26–45 | 45.8% | 43.2% | 2.5% | 43.6% | 2.2% |
| Age 46 or more | 16.9% | 14.2% | 2.8% | 15.8% | 1.2% |
| No-age data | 29.7% | 34.6% | 4.8% | 32.9% | 3.2% |
| Male | 41.1% | 37.4% | 3.7% | 38.9% | 2.2% |
| No-gender data | 29.8% | 34.6% | 4.8% | 33.0% | 3.2% |
| Product with discount | 77.0% | 79.8% | 2.8% | 77.7% | 0.8% |
| Segment online buyers | 7.6% | 6.4% | 1.2% | 7.0% | 0.6% |
| Segment others | 19.4% | 19.9% | 0.5% | 19.1% | 0.3% |
| Segment registered online | 13.6% | 13.1% | 0.5% | 12.8% | 0.8% |
| Segment registered offline only | 2.5% | 2.7% | 0.2% | 2.7% | 0.2% |
| Propensity score | 51.1% | 51.5% | 0.4% | 51.3% | 0.2% |
| Total | | | 2.1% | | 1.2% |

## Appendix D. Variance Inflation Factors

To evaluate if the correlation between explanatory variables can induce a multicollinearity problem in Table A4, we present the Variance Inflation Factors (VIF) for all explanatory variables we used for conversion and revenue equations. Here we include two specifications: The base model used in the article and the alternative specification we use as robustness in the appendices where we control for several additional covariates. In addition, for categorical variables, we report generalized inflation factors, GVIF (Fox & Monette, 1992).[4]

According to these results, we observe that for both specifications and both equation all VIFs are way below 5 or 10 that are some of the most common threshold used to detect problematic amount of collinearity (James et al., 2013, page 101).

Table A4
VIF for Conversion and Revenue Equations for two alternative specifications. For categorical variables, in addition of the individual VIF we include the generalized inflation factors (GVIF).

| | Conversion | | Revenue | |
|---|---|---|---|---|
| | Base | All covariates | Base | All covariates |
| Trigger | 1.006 | 1.008 | 1.007 | 1.010 |
| Narrow | 1.493 | 1.605 | 1.297 | 1.456 |
| Repeat | 1.851 | 1.897 | 1.807 | 1.843 |
| ShortD | 1.173 | 1.179 | 1.169 | 1.179 |
| ShortW | 1.589 | 1.597 | 1.494 | 1.508 |
| **Category (GVIF)** | 1.091 | 1.911 | 1.092 | 1.863 |
| Smartphone | 1.257 | 1.535 | 1.266 | 1.585 |
| Washing machine | 1.122 | 1.132 | 1.113 | 1.130 |
| Heater | 1.100 | 1.105 | 1.100 | 1.104 |
| Dryer | 1.122 | 1.815 | 1.122 | 1.748 |
| **Age (GVIF)** | | 2.470 | | 2.744 |
| 0–25 | | 1.567 | | 1.634 |
| 26–45 | | 3.107 | | 3.214 |
| >46 | | 2.224 | | 2.280 |
| **Segment (GVIF)** | | 2.336 | | 2.593 |
| Best online | | 2.051 | | 2.104 |
| Online buyers | | 1.465 | | 1.521 |
| Others | | 2.097 | | 2.166 |
| Registered | | 1.723 | | 1.871 |
| Open rate scores | | 1.017 | | 1.020 |
| Discounts | | 1.055 | | 1.056 |
| Banner | | 1.833 | | 1.848 |
| Morning | 1.413 | 1.423 | 1.454 | 1.468 |
| WCup | 1.917 | 1.945 | 1.829 | 1.846 |

## Appendix E. Robustness of Main Effects with Respect to Exclusion Restrictions and Estimation Approach

In our analysis of the effectiveness of triggered email marketing, we use a type-II Tobit model estimated through a 2-step method and using the World Cup and Morning dummies as exclusion restrictions. These choices were made to generate more precise and stable results, but the main insights do not critically depend on them. To demonstrate that the main effects are indeed robust, we estimated a series of 16 variations of the model. A graphic representation of the parameter estimates of the main effects of triggered emails for conversions and revenues are displayed in Fig. A5a and A5b respectively. To generate different models, we considered different alternatives for exclusion restrictions as well as estimation methods:

[4] Notice that standard routines to compute VIF require an intercept and therefore we need to discard one category intercept. In the results presented in Table A4, we discarded the baseline for LED TV.

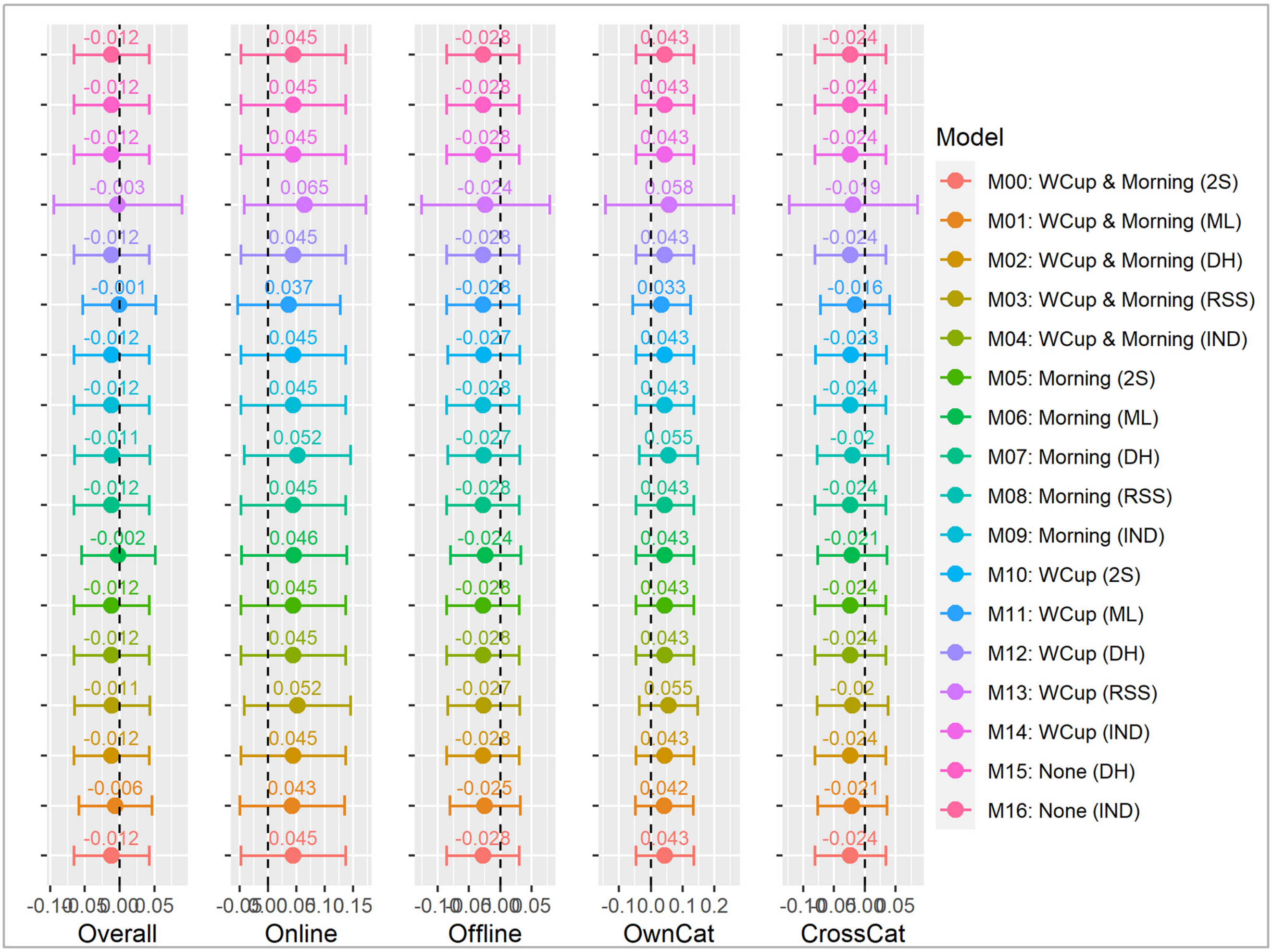


Fig. A5a. Effect of triggered emails in conversions for alternative specifications and estimation methods. Maximum likelihood and 2-step estimates are not reported for the case with no exclusion restrictions, because those fail to converge.

- Exclusion Restrictions. In addition to the use of World Cup and Morning dummies, we used in our base model, here we also explore variations where we only use the Morning dummy, only use the World Cup Dummy and others where we have no exclusion restrictions.
- Estimation Methods. Econometrics literature offers a number of alternative to estimate models with sample selection. In our analysis, we first considered a **Two-Step** (2S) method that was the original *control function* approach proposed by Heckman (1976). Second, we estimate a **Maximum Likelihood** approach (ML) that is statistically more efficient, but less numerically robust than the Two-Step method. Third, we use the **Double Hurdle** (DH) model proposed by Cragg (1971) and Blundell and Meghir, (1987) where we use a log- normal distribution for the error terms. Fourth, we consider a **Robust Sample Selection** (RSS) approach of Zhelonkin, Genton, and Ronchetti (2016) who proposed a robust estimate of the control function for a two-step estimation. Finally, we consider an **Independent** model approach (IND) where we assume there is no correlation in the error term between conversion and revenue equations (Duan, Manning, Morris, & Newhouse, 1984).

Results from Fig. A5a show that the main effects are very robust to variations in the estimation methods and the exclusion restriction. Regarding conversions, results indicate positive effects in online purchases in the targeted category, but the effects are not significant. Similarly, the negative coefficients in offline and cross-category suggest a potential substitution effect, but again those effects are not significant.

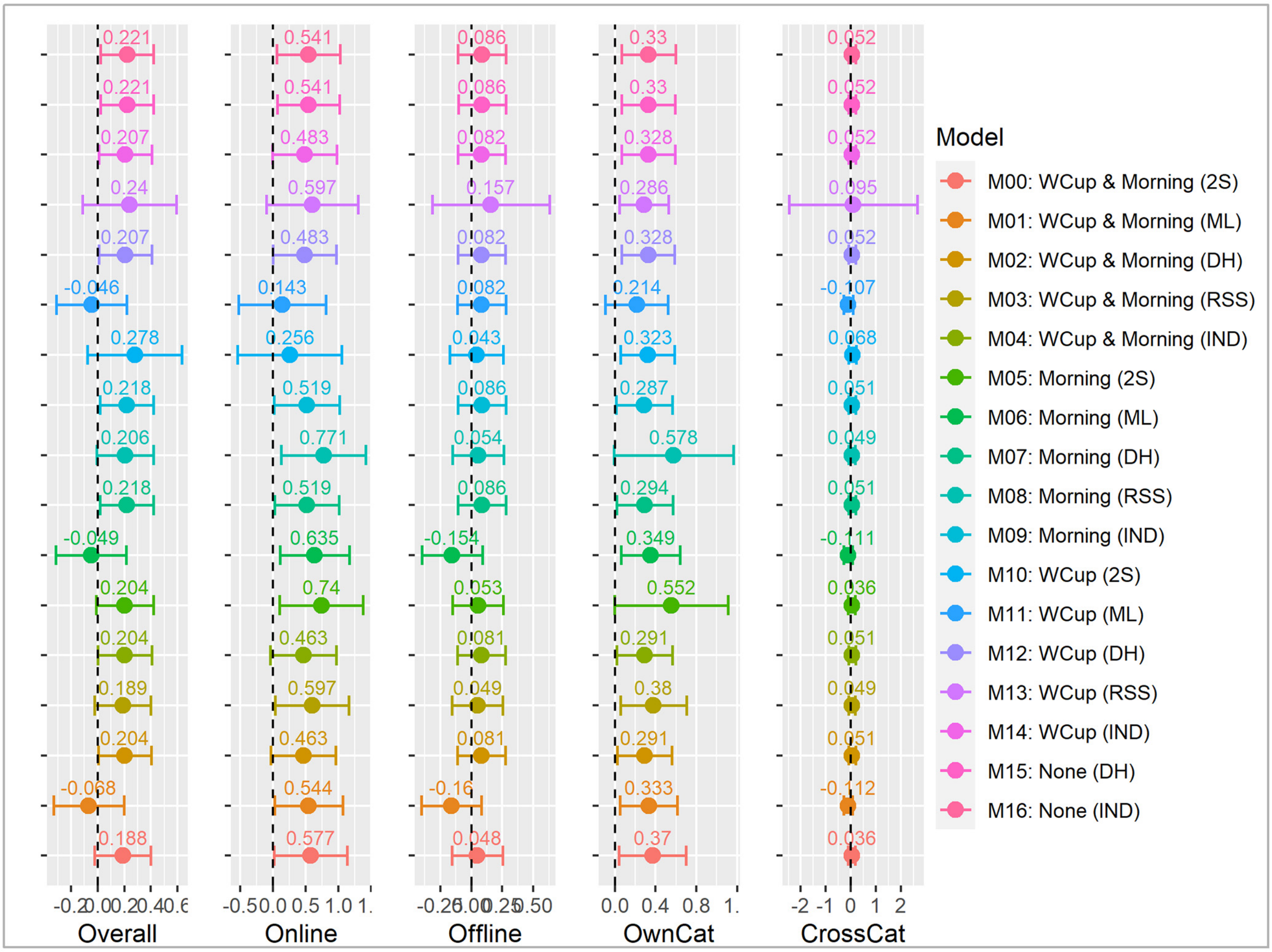


Fig. A5b. Effect of triggered emails in revenues for alternative specifications and estimation methods. Maximum likelihood and 2-step estimates are not reported for the case with no exclusion restrictions, because those fail to converge.

Results from Fig. A5b show that the main insights are also robust for the revenue equations. Here we find no consistent effect on offline and cross-category revenues, but a positive effect on online and targeted category. Notice that those models in which we only use the World Cup dummy as exclusion restrictions present larger dispersion providing additional justification to our choice of using two additional variables in the selection equation.

To complete this discussion, we note that the results of the conversion models appear to show less variability than the results of the revenue equations. This is expected because as the revenue equation is estimated conditional on converting, the conversion equation has much more data than the revenue equations. These results are, in addition, consistent with the alternative model specification of Appendices H and J, where we control for several other observables.

## Appendix F. Customer Characteristics Across Scenarios and Categories

Tables A6a show mean values for customer characteristics depending on the configuration of the triggered campaign showing no obvious differences between conditions.

Table A6a
Customer characteristics depending on the design of the triggered email.

| | Recommendation | | Repetition | | Delay | | Navigation window | |
|---|---|---|---|---|---|---|---|---|
| | Narrow | Broad | Yes | No | Short | Long | Short | Long |
| No age | 0.29 | 0.35 | 0.28 | 0.34 | 0.31 | 0.32 | 0.33 | 0.30 |
| Age 0–25 | 0.08 | 0.08 | 0.09 | 0.06 | 0.07 | 0.08 | 0.07 | 0.08 |
| Age 26–45 | 0.47 | 0.41 | 0.45 | 0.44 | 0.46 | 0.42 | 0.45 | 0.44 |
| Age 46 or more | 0.16 | 0.17 | 0.18 | 0.15 | 0.16 | 0.18 | 0.15 | 0.18 |
| No gender | 0.29 | 0.35 | 0.28 | 0.34 | 0.31 | 0.32 | 0.33 | 0.30 |
| F | 0.30 | 0.27 | 0.29 | 0.28 | 0.29 | 0.29 | 0.29 | 0.29 |
| M | 0.41 | 0.39 | 0.43 | 0.38 | 0.40 | 0.39 | 0.38 | 0.41 |
| No-segment | 0.41 | 0.45 | 0.40 | 0.45 | 0.42 | 0.43 | 0.44 | 0.41 |
| Best online | 0.18 | 0.18 | 0.18 | 0.17 | 0.18 | 0.17 | 0.17 | 0.18 |
| Online buyers | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.08 | 0.07 |
| Others | 0.20 | 0.18 | 0.20 | 0.19 | 0.19 | 0.20 | 0.19 | 0.20 |
| Registered | 0.14 | 0.12 | 0.15 | 0.12 | 0.13 | 0.13 | 0.12 | 0.14 |
| Opening rate | 50.3 | 54.3 | 51.6 | 53.6 | 52.7 | 52.8 | 52.8 | 52.7 |
| Discount | 0.77 | 0.78 | 0.79 | 0.76 | 0.76 | 0.80 | 0.77 | 0.78 |
| Banners | 0.87 | 0.71 | 0.79 | 0.82 | 0.82 | 0.77 | 0.80 | 0.81 |

Table A6b show mean values for customer characteristics depending on the target category, where some differences are expected. For example, interest for Heaters could be higher for more adult customers. Having said that, reported differences across demographic profiles do exist but are small.

Table A6b
Customer characteristics depending on the target product category.

| | LED | Smarthphone | Washing | Heater | Dryer |
|---|---|---|---|---|---|
| No age | 0.32 | 0.32 | 0.30 | 0.29 | 0.29 |
| Age 0–25 | 0.06 | 0.10 | 0.09 | 0.05 | 0.06 |
| Age 26–45 | 0.45 | 0.42 | 0.44 | 0.50 | 0.48 |
| Age 46 or more | 0.17 | 0.16 | 0.17 | 0.16 | 0.17 |
| No gender | 0.32 | 0.32 | 0.30 | 0.29 | 0.29 |
| Female | 0.23 | 0.30 | 0.39 | 0.40 | 0.33 |
| Male | 0.45 | 0.38 | 0.31 | 0.31 | 0.38 |
| No-segment | 0.42 | 0.44 | 0.44 | 0.45 | 0.38 |
| Best online | 0.18 | 0.17 | 0.17 | 0.17 | 0.20 |
| Online buyers | 0.07 | 0.07 | 0.08 | 0.08 | 0.09 |
| Others | 0.19 | 0.18 | 0.20 | 0.20 | 0.22 |
| Registered | 0.15 | 0.13 | 0.12 | 0.10 | 0.11 |
| Opening rate | 53.6 | 50.9 | 52.9 | 54.5 | 53.7 |
| Discount | 0.77 | 0.81 | 0.75 | 0.74 | 0.74 |
| Banners | 1.00 | 0.67 | 0.89 | 1.00 | 0.24 |

## Appendix G. Effect by Department/Category

In the main model we only use category dummies to define the corresponding baselines, but we did not interact with the treatment to see if triggered emails are more effective in some categories than others. Table A7a reports the results of the models with those interactions. As expected, in this more granular model only a small number of coefficients are significant. For example, results suggest that triggered email could be particularly detrimental in offline conversions of heaters and dryers and that triggered emails are particularly effective in stimulating revenues for LED TVs.

Table A7a
Tobit regression coefficients, for the average treatment effects with interactions by category.

| | Total | | Channel Decomposition | | | | Category Decomposition | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Overall | p-val | Online | p-val | Offline | p-val | OwnCat | p-val | CrossCat | p-val |
| *Conversion equation* | | | | | | | | | | |
| LED | −1.561 | <0.001 | −2.310 | <0.001 | −1.646 | <0.001 | −2.314 | <0.001 | −1.648 | <0.001 |
| Smartphone | −1.643 | <0.001 | −2.338 | <0.001 | −1.750 | <0.001 | −2.197 | <0.001 | −1.778 | <0.001 |
| Washing | −1.477 | <0.001 | −2.261 | <0.001 | −1.567 | <0.001 | −2.176 | <0.001 | −1.587 | <0.001 |
| Heater | −1.413 | <0.001 | −2.458 | <0.001 | −1.440 | <0.001 | −2.439 | <0.001 | −1.442 | <0.001 |
| Dryer | −1.433 | <0.001 | −2.329 | <0.001 | −1.486 | <0.001 | −2.299 | <0.001 | −1.489 | <0.001 |
| Trigger x LED | 0.039 | 0.440 | 0.008 | 0.926 | 0.037 | 0.489 | 0.101 | 0.240 | 0.022 | 0.687 |
| Trigger x Smartphone | −0.020 | 0.734 | 0.050 | 0.596 | −0.034 | 0.588 | 0.019 | 0.831 | −0.043 | 0.506 |
| Trigger x Washing | 0.077 | 0.466 | 0.237 | 0.167 | 0.026 | 0.815 | 0.167 | 0.323 | 0.076 | 0.499 |
| Trigger x Heater | −0.179 | 0.148 | 0.090 | 0.715 | **−0.238** | **0.066** | −0.172 | 0.537 | −0.183 | 0.150 |
| Trigger x Dryer | −0.163 | 0.113 | −0.066 | 0.730 | **−0.178** | **0.097** | −0.234 | 0.264 | −0.131 | 0.215 |
| Morning | 0.286 | <0.001 | 0.251 | <0.001 | 0.277 | <0.001 | 0.132 | 0.020 | 0.298 | <0.001 |
| WCup | 0.042 | 0.202 | 0.100 | 0.075 | 0.011 | 0.746 | 0.149 | 0.007 | −0.001 | 0.969 |
| *Revenue equation* | | | | | | | | | | |
| LED | −0.130 | 0.898 | −3.948 | 0.266 | −0.243 | 0.818 | −0.315 | 0.909 | 0.628 | 0.364 |
| Smartphone | −0.010 | 0.992 | −4.091 | 0.249 | −0.058 | 0.958 | −0.351 | 0.895 | 0.463 | 0.519 |
| Washing machine | 0.177 | 0.861 | −2.582 | 0.473 | −0.174 | 0.868 | 0.297 | 0.911 | 0.736 | 0.290 |
| Heater | −0.329 | 0.746 | −4.690 | 0.231 | −0.208 | 0.839 | −0.755 | 0.801 | 0.698 | 0.299 |
| Dryer | −0.378 | 0.707 | −4.644 | 0.200 | −0.299 | 0.772 | −1.764 | 0.526 | 0.666 | 0.327 |
| Trigger x LED | 0.219 | 0.269 | 0.841 | 0.108 | 0.103 | 0.584 | **0.694** | **0.037** | −0.079 | 0.551 |
| Trigger x Smartphone | 0.163 | 0.483 | 0.629 | 0.266 | −0.049 | 0.830 | 0.226 | 0.480 | 0.069 | 0.672 |
| Trigger x Washing | 0.335 | 0.405 | 0.107 | 0.918 | 0.123 | 0.750 | 0.215 | 0.726 | 0.012 | 0.965 |
| Trigger x Heater | 0.260 | 0.597 | 0.841 | 0.583 | 0.084 | 0.856 | −0.196 | 0.855 | 0.471 | 0.132 |
| Trigger x Dryer | −0.055 | 0.892 | −0.404 | 0.729 | −0.009 | 0.982 | −0.001 | 0.999 | 0.104 | 0.685 |
| *Joint parameters* | | | | | | | | | | |
| Inv. Mills Ratio | 1.585 | 0.004 | 3.227 | 0.024 | 1.380 | 0.011 | 2.426 | 0.027 | 0.767 | 0.031 |
| σ | 2.406 | | 3.709 | | 2.090 | | 2.573 | | 1.382 | |
| *P* | 0.659 | | 0.870 | | 0.660 | | 0.943 | | 0.555 | |

To understand if the differences in the interactions are significant, we need to evaluate a series of pairwise evaluations. Table A7b and A7c report the mean differences and standard deviations for the selection and outcome respectively considering total sales (equivalent matrices by channel and category are available upon request). Results from Tables A7b and A7c show no evidence of systematic differences in the effect of the treatment depending on the category. These results justify our decision of not considering interactions in the main model.

Table A7b
Difference in interaction coefficients for all pairs of categories for the selection equation (standard error in parenthesis).

| | LED | Smartphone | Washing | Heater |
|---|---|---|---|---|
| Smartphone | −0.059 (0.077) | – | – | – |
| Washing | 0.038 (0.117) | 0.097 (0.121) | – | – |
| Heater | −0.218 (0.134) | −0.159 (0.137) | −0.256 (0.163) | – |
| Dryer | −0.202 (0.115) | −0.144 (0.118) | −0.241 (0.148) | 0.016 (0.161) |

Table A7c
Difference in interaction coefficients for all pairs of categories for the output equation (standard error in parenthesis).

| | LED | Smartphone | Washing | Heater |
|---|---|---|---|---|
| Smartphone | −0.056 (0.305) | – | – | – |
| Washing | 0.116 (0.448) | 0.172 (0.464) | – | – |
| Heater | 0.041 (0.530) | 0.097 (0.544) | −0.075 (0.635) | – |
| Dryer | −0.274 (0.452) | −0.218 (0.469) | −0.390 (0.572) | −0.315 (0.639) |

## Appendix H. Estimation of Marginal Effects

As is pointed out by Vance (2009), the marginal effect must be calculated using a nonlinear function of the underlying model parameters to correct for the selectivity effect. In our case, the triggered treatment affects the selection and output equations and therefore we estimate the effect on the purchase probability, the effect on revenues conditional on purchase and the unconditional effect on revenues, as reported in Table A8.

Table A8
Marginal effects.

| | Purchase probability | | Conditional revenue | | Unconditional revenue | |
|---|---|---|---|---|---|---|
| | Mean | s.d | Mean | s.d | Mean | s.d |
| Online | 0.002 | 0.001 | 14.866 | 7.276 | 0.791 | 0.369 |
| Offline | −0.004 | 0.001 | 0.388 | 0.080 | 0.008 | 0.002 |
| Own Category | 0.002 | 0.001 | 36.786 | 13.938 | 0.860 | 0.336 |
| Cross Category | −0.003 | 0.001 | 0.159 | 0.020 | 0.001 | 0.001 |
| Total | −0.002 | 0.000 | 1.474 | 0.370 | 0.102 | 0.021 |

The exact computation of the first two columns is direct. The marginal effect in the selection equation, is equivalent to the marginal effect in a probit model (see for example, Fernihough, 2011). For the case of the output equation, we use the linear estimate corrected by the effect of the selection (Saha et al., 1997)

$$\Delta_T \ \Pr(y_i = 1) = \Phi\left(\hat{\alpha} + \hat{\theta}_1' \mathbf{x}_{1i}\right) - \Phi\left(\hat{\theta}_1' \mathbf{x}_{1i}\right) \tag{A1}$$

$$\Delta_T \mathrm{E}(\ \ln(r_i)|y_i = 1) = \hat{\beta} + \hat{\lambda}\left[\frac{\phi\left(\hat{\alpha} + \hat{\theta}_1' \mathbf{x}_{1i}\right)}{\Phi\left(\hat{\alpha} + \hat{\theta}_1' \mathbf{x}_{1i}\right)} - \frac{\phi\left(\hat{\theta}_1' \mathbf{x}_{1i}\right)}{\Phi\left(\hat{\theta}_1' \mathbf{x}_{1i}\right)}\right] \tag{A2}$$

In eq. (A2), $\hat{\lambda}$ represent the estimate of the coefficient multiplying the Inverse Mills Ratio. To compute the unconditional effect on revenues, we follow McDonald and Moffitt (1980) to argue that change in the (log) revenues can be decomposed in the change in the expected value of the revenues among the positive revenues, weighted by the probability of being positive (given a treatment) and the change in the probability of being having positive revenues, weighted by the conditional expected revenue (given no treatment)

$$\Delta_T \mathrm{E}(\ln(r_i)) = P_o(y_i = 1)[E_1(\ln(r_i)|y_i = 1) - E_0(\ln(r_i)|y_i = 1)] + E_1(\ln(r_i)|y_i = 1)[P_1(y_i = 1) - P_o(y_i = 1)] \tag{A3}$$

All these marginal calculations depend on $\mathbf{x}_{1i}$ and therefore we compute them for each individual. The values reported in Table A8 are the mean across all individuals. Eqs. (A2) and (A3) provides an estimate of the marginal effect in the log scale, but estimates in Table A8 are already exponentiated to represent monetary values. The intuition of the marginal effects is straightforward. Consider for example the case of online revenues. Triggered emails have a very small impact on conversions, but conditional on a purchase, the intervention increases revenues by US$14.9. Considering only a small fraction of customer purchase, the unconditional effect per email send is US$0.791. These averages consider all configurations including those that are less effective. Firms interested in implementing triggered emails are likely to adopt those configurations that show more promising results. For example, according to our estimates, the counterfactual scenarios where all emails were sent using narrow recommendations would imply an online revenue of US$1.12 and if the emails were sent with a short delay, the online revenues per email would be as high as US$1.74.

## Appendix I. Alternative Specifications for the Average Effect Regression

Table A9
Tobit regression coefficients, for the average treatment effects with additional controls.

| | Total | | Channel Decomposition | | | | Category Decomposition | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Purchase | p-val | Online | p-val | Offline | p-val | Own | p-val | Cross | p-val |
| *Conversion equation* | | | | | | | | | | |
| Trigger | −0.009 | 0.792 | 0.047 | 0.400 | −0.026 | 0.455 | 0.040 | 0.467 | −0.020 | 0.563 |
| LED | −1.697 | <0.001 | −2.140 | <0.001 | −1.869 | <0.001 | −2.134 | <0.001 | −1.875 | <0.001 |
| Smartphone | −1.793 | <0.001 | −2.214 | <0.001 | −1.976 | <0.001 | −2.121 | <0.001 | −2.001 | <0.001 |
| Washing | −1.587 | <0.001 | −1.981 | <0.001 | −1.786 | <0.001 | −1.970 | <0.001 | −1.775 | <0.001 |
| Heater | −1.673 | <0.001 | −2.239 | <0.001 | −1.817 | <0.001 | −2.386 | <0.001 | −1.788 | <0.001 |
| Dryer | −1.650 | <0.001 | −2.401 | <0.001 | −1.738 | <0.001 | −2.429 | <0.001 | −1.712 | <0.001 |
| Age 0–25 | 0.083 | 0.302 | −0.351 | 0.028 | 0.201 | 0.016 | −0.252 | 0.088 | 0.185 | 0.028 |
| Age 26–45 | 0.238 | <0.001 | −0.224 | 0.057 | 0.333 | <0.001 | −0.078 | 0.450 | 0.297 | <0.001 |
| Age 46 or more | 0.147 | 0.026 | −0.155 | 0.221 | 0.192 | 0.006 | 0.014 | 0.903 | 0.158 | 0.026 |
| Best online | −0.014 | 0.817 | 0.417 | <0.001 | −0.105 | 0.091 | 0.085 | 0.435 | −0.036 | 0.567 |
| Online buyers | −0.099 | 0.197 | 0.384 | 0.005 | −0.224 | 0.007 | 0.111 | 0.400 | −0.141 | 0.082 |
| Others | −0.024 | 0.679 | 0.280 | 0.019 | −0.087 | 0.152 | 0.124 | 0.242 | −0.049 | 0.421 |
| Registered | −0.167 | 0.011 | 0.018 | 0.901 | −0.196 | 0.004 | 0.074 | 0.524 | −0.213 | 0.002 |
| Open rate score | 0.000 | 0.949 | 0.001 | 0.398 | 0.000 | 0.666 | 0.000 | 0.953 | 0.000 | 0.754 |
| Discount | 0.027 | 0.508 | −0.083 | 0.208 | 0.046 | 0.281 | 0.020 | 0.774 | 0.018 | 0.668 |
| Banner | 0.031 | 0.572 | −0.224 | 0.012 | 0.105 | 0.077 | −0.182 | 0.032 | 0.114 | 0.058 |
| Morning | 0.284 | <0.001 | 0.285 | <0.001 | 0.268 | <0.001 | 0.154 | 0.007 | 0.293 | <0.001 |
| World Cup | 0.053 | 0.114 | 0.095 | 0.094 | 0.028 | 0.440 | 0.146 | 0.008 | 0.014 | 0.705 |
| *Revenue equation* | | | | | | | | | | |
| Trigger | 0.145 | 0.134 | 0.423 | 0.103 | 0.028 | 0.759 | 0.387 | 0.042 | 0.022 | 0.688 |
| LED | 1.534 | 0.069 | 0.173 | 0.943 | 1.059 | 0.234 | 0.405 | 0.862 | 1.887 | <0.001 |
| Smartphone | 1.403 | 0.100 | −0.566 | 0.817 | 1.064 | 0.239 | −0.104 | 0.964 | 1.768 | 0.001 |
| Washing | 1.799 | 0.029 | 0.711 | 0.763 | 1.108 | 0.204 | 0.673 | 0.763 | 1.975 | <0.001 |
| Heater | 1.326 | 0.127 | −0.257 | 0.922 | 0.982 | 0.275 | −0.471 | 0.857 | 2.079 | <0.001 |
| Dryer | 0.759 | 0.363 | −2.026 | 0.439 | 0.801 | 0.348 | −1.841 | 0.472 | 1.891 | <0.001 |
| Age 0–25 | −0.542 | 0.023 | −2.115 | 0.007 | −0.101 | 0.653 | −1.463 | 0.009 | 0.038 | 0.778 |
| Age 26–45 | −0.190 | 0.303 | −1.085 | 0.051 | 0.146 | 0.441 | −0.450 | 0.202 | 0.088 | 0.417 |
| Age 46 or more | −0.048 | 0.810 | −0.249 | 0.659 | 0.012 | 0.949 | −0.251 | 0.506 | −0.012 | 0.917 |
| Best online | 0.288 | 0.088 | 1.279 | 0.039 | −0.157 | 0.321 | 0.643 | 0.080 | 0.144 | 0.125 |
| Online buyers | 0.154 | 0.495 | 0.957 | 0.154 | −0.189 | 0.401 | 0.399 | 0.378 | −0.087 | 0.499 |
| Others | 0.214 | 0.200 | 0.879 | 0.128 | −0.059 | 0.700 | 0.333 | 0.362 | −0.001 | 0.988 |
| Registered | 0.063 | 0.755 | 0.437 | 0.500 | 0.006 | 0.975 | 0.399 | 0.310 | −0.147 | 0.206 |
| Open rate score | 0.000 | 0.992 | −0.006 | 0.288 | 0.001 | 0.623 | 0.000 | 0.955 | 0.000 | 0.831 |
| Discount | 0.088 | 0.459 | 0.037 | 0.902 | 0.199 | 0.075 | 0.195 | 0.404 | −0.002 | 0.982 |
| Banner | −0.513 | 0.002 | −1.265 | 0.005 | −0.163 | 0.308 | −0.614 | 0.053 | −0.107 | 0.277 |
| *Joint parameter* | | | | | | | | | | |
| Inv. Mills Ratio | 1.364 | 0.001 | 2.464 | 0.014 | 1.140 | 0.005 | 2.436 | 0.011 | 0.591 | 0.010 |
| $\sigma$ | 1.882 | | 2.805 | | 1.580 | | 2.516 | | 0.913 | |
| $\rho$ | 0.725 | | 0.878 | | 0.721 | | 0.968 | | 0.647 | |

Results are similar to those presented in Table 6, being the most notable difference that now the effect of triggers in online revenues is barely significant. Notice however that the interpretation of the main effects depends on the codification of the other dummies we use. For example, the effect of triggered email reported on the table correspond to customers with no Age information, which is the level we excluded.

## Appendix J. Alternative Specifications for the Campaign Design Regression

Table A10
Tobit regression coefficients, for the campaign design effects with additional controls.

| | Total | | Channel decomposition | | | | Category decomposition | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Sales | p-val | Online | p-val | Offline | p-val | OwnCat | p-val | CrossCat | p-val |
| *Conversion equation* | | | | | | | | | | |
| Triggger | −0.006 | 0.860 | 0.044 | 0.438 | −0.022 | 0.532 | 0.036 | 0.520 | −0.016 | 0.644 |
| Narrow | 0.195 | <0.001 | 0.060 | 0.420 | 0.228 | 0.000 | −0.121 | 0.081 | 0.297 | <0.001 |
| Repeat | 0.034 | 0.459 | 0.245 | 0.002 | −0.049 | 0.311 | 0.147 | 0.055 | −0.016 | 0.749 |
| ShortD | −0.040 | 0.317 | −0.044 | 0.525 | −0.023 | 0.596 | 0.142 | 0.040 | −0.094 | 0.030 |
| ShortW | −0.290 | <0.001 | −0.156 | 0.035 | −0.318 | <0.001 | −0.069 | 0.316 | −0.355 | <0.001 |
| LED | −1.637 | <0.001 | −2.278 | <0.001 | −1.764 | <0.001 | −2.302 | <0.001 | −1.769 | <0.001 |
| Smartphone | −1.715 | <0.001 | −2.297 | <0.001 | −1.863 | <0.001 | −2.256 | <0.001 | −1.883 | <0.001 |
| Washing | −1.529 | <0.001 | −2.107 | <0.001 | −1.688 | <0.001 | −2.134 | <0.001 | −1.674 | <0.001 |
| Heater | −1.623 | <0.001 | −2.380 | <0.001 | −1.722 | <0.001 | −2.530 | <0.001 | −1.697 | <0.001 |
| Dryer | −1.606 | <0.001 | −2.469 | <0.001 | −1.666 | <0.001 | −2.534 | <0.001 | −1.640 | <0.001 |
| Age 0–25 | 0.074 | 0.361 | −0.372 | 0.021 | 0.196 | 0.020 | −0.272 | 0.068 | 0.182 | 0.032 |
| Age 26–45 | 0.228 | <0.001 | −0.220 | 0.062 | 0.321 | <0.001 | −0.077 | 0.458 | 0.286 | <0.001 |
| Age 46 or more | 0.128 | 0.053 | −0.174 | 0.170 | 0.174 | 0.014 | 0.005 | 0.962 | 0.138 | 0.053 |
| Best online | −0.012 | 0.839 | 0.417 | 0.000 | −0.103 | 0.099 | 0.082 | 0.450 | −0.032 | 0.609 |
| Online buyers | −0.092 | 0.233 | 0.386 | 0.005 | −0.216 | 0.009 | 0.118 | 0.376 | −0.134 | 0.103 |
| Others | −0.027 | 0.645 | 0.276 | 0.022 | −0.089 | 0.145 | 0.127 | 0.236 | −0.053 | 0.390 |
| Registered | −0.170 | 0.010 | 0.010 | 0.943 | −0.199 | 0.004 | 0.078 | 0.504 | −0.219 | 0.002 |
| Open rate score | 0.000 | 0.658 | 0.001 | 0.400 | −0.001 | 0.403 | 0.000 | 0.881 | −0.001 | 0.432 |
| Discount | 0.024 | 0.552 | −0.098 | 0.142 | 0.048 | 0.272 | 0.023 | 0.742 | 0.014 | 0.741 |
| Banner | 0.004 | 0.951 | −0.154 | 0.107 | 0.056 | 0.376 | −0.111 | 0.226 | 0.060 | 0.351 |
| Morning | 0.143 | <0.001 | 0.139 | 0.046 | 0.139 | 0.001 | 0.131 | 0.049 | 0.128 | 0.002 |
| WCup | 0.178 | <0.001 | 0.274 | 0.001 | 0.119 | 0.016 | 0.170 | 0.028 | 0.175 | 0.001 |
| *Revenue equation* | | | | | | | | | | |
| Trigger | 0.119 | 0.227 | 0.339 | 0.140 | 0.008 | 0.935 | 0.388 | 0.055 | 0.005 | 0.928 |
| Narrow | −0.620 | <0.001 | −1.295 | <0.001 | −0.241 | 0.158 | −1.228 | <0.001 | −0.034 | 0.749 |
| Repeat | 0.294 | 0.016 | 0.288 | 0.321 | 0.048 | 0.726 | 0.332 | 0.142 | 0.026 | 0.751 |
| ShortD | 0.286 | 0.016 | 0.627 | 0.030 | 0.123 | 0.297 | 0.685 | 0.011 | −0.051 | 0.509 |
| ShortW | 0.139 | 0.512 | 0.780 | 0.020 | −0.224 | 0.388 | 0.187 | 0.487 | −0.060 | 0.695 |
| LED | 0.478 | 0.671 | −0.129 | 0.957 | −0.509 | 0.727 | −1.923 | 0.501 | 0.924 | 0.236 |
| Smartphone | 0.447 | 0.697 | −0.597 | 0.804 | −0.486 | 0.746 | −2.141 | 0.441 | 0.790 | 0.328 |
| Washing | 0.884 | 0.419 | 0.440 | 0.849 | −0.331 | 0.817 | −1.494 | 0.583 | 1.084 | 0.155 |
| Heater | 0.425 | 0.708 | −0.459 | 0.856 | −0.460 | 0.751 | −2.563 | 0.413 | 1.168 | 0.127 |
| Dryer | −0.019 | 0.987 | −1.814 | 0.479 | −0.561 | 0.690 | −3.835 | 0.211 | 1.018 | 0.171 |
| Age 0–25 | −0.565 | 0.020 | −1.848 | 0.013 | −0.008 | 0.976 | −1.718 | 0.005 | 0.086 | 0.576 |
| Age 26–45 | −0.077 | 0.698 | −0.792 | 0.121 | 0.355 | 0.157 | −0.473 | 0.213 | 0.205 | 0.128 |
| Age 46 or more | −0.013 | 0.950 | −0.262 | 0.609 | 0.120 | 0.588 | −0.275 | 0.497 | 0.033 | 0.799 |
| Best online | 0.270 | 0.118 | 1.097 | 0.064 | −0.221 | 0.223 | 0.703 | 0.076 | 0.136 | 0.193 |
| Online buyers | 0.141 | 0.541 | 0.843 | 0.187 | −0.306 | 0.250 | 0.603 | 0.220 | −0.136 | 0.348 |
| Others | 0.223 | 0.190 | 0.857 | 0.105 | −0.098 | 0.574 | 0.500 | 0.208 | −0.012 | 0.905 |
| Registered | 0.072 | 0.737 | 0.322 | 0.578 | −0.073 | 0.747 | 0.525 | 0.215 | −0.220 | 0.116 |
| Open rate score | 0.001 | 0.761 | 0.001 | 0.904 | 0.001 | 0.785 | 0.004 | 0.365 | 0.000 | 0.991 |
| Discount | 0.145 | 0.237 | 0.143 | 0.599 | 0.256 | 0.044 | 0.324 | 0.197 | 0.020 | 0.784 |
| Banner | −0.277 | 0.113 | −0.688 | 0.100 | −0.029 | 0.874 | −0.169 | 0.623 | −0.064 | 0.565 |
| ***Joint parameters*** | 1.736 | 0.003 | 2.148 | 0.034 | 1.890 | 0.008 | 3.008 | 0.008 | 1.083 | 0.005 |
| Inv. Mills Ratio | 2.090 | #N/D | 2.458 | #N/D | 2.088 | #N/D | 2.989 | #N/D | 1.229 | #N/D |
| σ | 0.831 | #N/D | 0.874 | #N/D | 0.905 | #N/D | 1.006 | #N/D | 0.881 | #N/D |
| ρ | 0.119 | 0.227 | 0.339 | 0.140 | 0.008 | 0.935 | 0.388 | 0.055 | 0.005 | 0.928 |

These results show that, after controlling for all observables, we have access to the significance and sign of all coefficients associated with the design of triggered emails are identical to those reported in Table 7. We conclude that the analysis presented in the main model are robust to alternative model specifications.

## Appendix K. Sequential Analysis of Absolute Effect of Design Components

In the main model presented in Table 6, we jointly analyze the marginal effect of different configurations. In Table 7, the reported estimates correspond to the change in conversion and revenues with respect to the base configurations (Broad recommendations, No repetition, Long Delay and Long navigation window). To evaluate if a given level is better than the non-treatment alternative, we run a sequence of models when we compare one factor at the time as reported in Table A11. In this table, parameter estimates correspond to the marginal effect with respect to the non-treatment that we leave as the baseline.

Table A11
Effect of different levels of design with respect to non-treatment.

| | (M1) | | (M2) | | (M3) | | (M4) | |
|---|---|---|---|---|---|---|---|---|
| | Estimate | p-val | Estimate | p-val | Estimate | p-val | Estimate | p-val |
| *Conversion equation* | | | | | | | | |
| LED | −1.536 | 0.000 | −1.541 | 0.000 | −1.535 | 0.000 | −1.519 | 0.000 |
| Smartphone | −1.646 | 0.000 | −1.649 | 0.000 | −1.646 | 0.000 | −1.620 | 0.000 |
| Washing | −1.430 | 0.000 | −1.433 | 0.000 | −1.429 | 0.000 | −1.414 | 0.000 |
| Heater | −1.502 | 0.000 | −1.503 | 0.000 | −1.498 | 0.000 | −1.474 | 0.000 |
| Dryer | −1.507 | 0.000 | −1.510 | 0.000 | −1.506 | 0.000 | −1.486 | 0.000 |
| CON.Morning | 0.281 | 0.000 | 0.276 | 0.000 | 0.286 | 0.000 | 0.242 | 0.000 |
| CON.WCup | 0.048 | 0.159 | 0.062 | 0.087 | 0.040 | 0.234 | 0.049 | 0.139 |
| Trigger x Narrow | 0.003 | 0.931 | | | | | | |
| Trigger x Broad | −0.035 | 0.436 | | | | | | |
| Trigger x Repeat | | | 0.025 | 0.554 | | | | |
| Trigger x NoRepeat | | | −0.046 | 0.265 | | | | |
| Trigger x ShortD | | | | | −0.006 | 0.866 | | |
| Trigger x LongD | | | | | −0.026 | 0.605 | | |
| Trigger x ShortW | | | | | | | −0.157 | 0.001 |
| Trigger x LongW | | | | | | | 0.073 | 0.049 |
| *Revenue equation* | | | | | | | | |
| LED | −0.123 | 0.902 | 0.045 | 0.965 | 0.115 | 0.908 | −0.015 | 0.990 |
| Smartphone | −0.003 | 0.998 | 0.146 | 0.889 | 0.208 | 0.839 | 0.093 | 0.941 |
| Washing | 0.296 | 0.762 | 0.413 | 0.677 | 0.467 | 0.631 | 0.375 | 0.750 |
| Heater | −0.186 | 0.858 | −0.143 | 0.891 | −0.075 | 0.942 | −0.204 | 0.867 |
| Dryer | −0.451 | 0.658 | −0.330 | 0.749 | −0.264 | 0.794 | −0.384 | 0.752 |
| Trigger x Narrow | −0.088 | 0.541 | | | | | | |
| Trigger x Broad | 0.664 | 0.000 | | | | | | |
| Trigger x Repeat | | | 0.165 | 0.284 | | | | |
| Trigger x NoRepeat | | | 0.219 | 0.175 | | | | |
| Trigger x ShortD | | | | | 0.316 | 0.024 | | |
| Trigger x LongD | | | | | −0.120 | 0.524 | | |
| Trigger x ShortW | | | | | | | 0.214 | 0.342 |
| Trigger x LongW | | | | | | | 0.204 | 0.178 |
| *Joint parameters* | | | | | | | | |
| Inv. Mills Ratio | 1.574 | 0.004 | 1.495 | 0.007 | 1.460 | 0.007 | 1.525 | 0.021 |
| $\sigma$ | 2.385 | | 2.358 | | 2.335 | | 2.372 | |
| $\rho$ | 0.660 | | 0.634 | | 0.625 | | 0.643 | |

Results in Table A11 correspond to total conversion and revenues. Equivalent tables with effects by channel and category are available upon request. In model 1 (M1), we analyze the effect of different recommendations content (Narrow vs. Broad Assortment). In model 2 (M2), we analyze the effect of repetition (Repeat vs. No Repetition). In model 3 (M3), we analyze the effect of the delay from the identification of a browsing event and the delivery of the email (Short vs. Long). Finally, model 4 (M4) analyzes the effect of the length of navigation windows used to decide which customer are eligible to receive a triggered email (Short vs. Long navigation window).

Results of Table A11 reveal that in some cases the reported effects of triggers are only driven by one level. For example, when we look at the delay from the identification of the triggered event to the delivery of the message, we find that short delays generate positive and significant impact, but waiting longer does not lead to larger revenues.

## References


Amemiya, T. (1984). Tobit models: a survey. *Journal of Econometrics, 24*(1–2), 3–61.

Anderson, S., & Newell, R. G. (2003). Simplified marginal effects in discrete choice models. *Economics Letters, 81*(3), 321–326.

Ansari, A., & Mela, C. F. (2003). E-customization. *Journal of Marketing Research, 40*(2), 131–145.

Arora, N., Dreze, X., Ghose, A., Hess, J. D., Iyengar, R., & Jing, B., et al (2008). Putting one-to-one marketing to work: Personalization, customization, and choice. *Marketing Letters, 19*(3–4), 305–321.

Blattberg, R. C., Kim, B.-D., & Neslin, S. A. (2008). Customer privacy and database marketing. *Database Marketing: Analyzing and Managing Customers*, 75–101.

Blundell, R., & Meghir, C. (1987). Bivariate alternatives to the Tobit model. *Journal of Econometrics, 34*(1–2), 179–200.

Bonfrer, A., & Drèze, X. (2009). Real-time evaluation of e-mail campaign performance. *Marketing Science, 28*(2), 251–263.

Breugelmans, E., & Campo, K. (2016). Cross-channel effects of price promotions: an empirical analysis of the multi-channel grocery retail sector. *Journal of Retailing, 92*(3), 333–351.

Bucklin, R. E., & Sismeiro, C. (2003). A model of web site browsing behavior estimated on clickstream data. *Journal of Marketing Research, 40*(3), 249–267.

Bult, J. R., & Wansbeek, T. (1995). Optimal selection for direct mail. *Marketing Science, 14*(4), 378–394.

Chandon, P., Wansink, B., & Laurent, G. (2000). A benefit congruency framework of sales promotion effectiveness. *Journal of Marketing, 64*(4), 65–81.

Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica*, 829–844.

Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and Statistics, 84*(1), 151–161.

Dinner, I. M., Van Heerde, H. J., & Neslin, S. A. (2014). Driving online and offline sales: The cross-channel effects of traditional, online display, and paid search advertising. *Journal of Marketing Research, 51*(5), 527–545.

Duan, N., Manning, W. G., Morris, C. N., & Newhouse, J. P. (1984). Choosing between the sample-selection model and the multi-part model. *Journal of Business & Economic Statistics, 2*(3), 283–289.

Feng, C., & Fay, S. (2020). Store Closings and Retailer profitability: A Contingency Perspective. *Journal of Retailing, 96*(3), 411–433.

Fernihough, A. (2011). *Simple logit and probit marginal effects in R*.

Fox, E. J., Montgomery, A. L., & Lodish, L. M. (2004). Consumer shopping and spending across retail formats. *The Journal of Business, 77*(S2), S25–S60.

Fox, J., & Monette, G. (1992). Generalized collinearity diagnostics. *Journal of the American Statistical Association, 87*(417), 178–183.

Gallino, S., & Moreno, A. (2014). Integration of online and offline channels in retail: The impact of sharing reliable inventory availability information. *Management Science, 60*(6), 1434–1451.

Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation*. WW Norton.

Goic, M., Álvarez, R., & Montoya, R. (2018). The effect of house ads on multichannel sales. *Journal of Interactive Marketing, 42*, 32–45.

Goic, M., & Olivares, M. (2019). Omnichannel analytics. *Operations in an Omnichannel World*. Cham: Springer, 115–150.

Golrezaei, N., Nazerzadeh, H., & Rusmevichientong, P. (2014). Real-time optimization of personalized assortments. *Management Science, 60*(6), 1532–1551.

Gordon, B. R., Zettelmeyer, F., Bhargava, N., & Chapsky, D. (2019). A comparison of approaches to advertising measurement: Evidence from big field experiments at Facebook. *Marketing Science, 38*(2), 193–225.

Gubela, R. M., Lessmann, S., & Jaroszewicz, S. (2020). Response transformation and profit decomposition for revenue uplift modeling. *European Journal of Operational Research, 283*(2), 647–661.

Gunst, R. F., & Mason, R. L. (2009). Fractional factorial design. *Wiley Interdisciplinary Reviews: Computational Statistics, 1*(2), 234–244.

Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of economic and social measurement, NBER, 475–492.*

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica, 47*(1), 153–161.

Hertwig, R. (2012). Tapping into the wisdom of the crowd—with confidence. *Science, 336*(6079), 303–304.

Hoch, S. J., Bradlow, E. T., & Wansink, B. (1999). The variety of an assortment. *Marketing Science, 18*(4), 527–546.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning,* , Vol. 112. New York: Springer18.

Kollmann, T., Kuckertz, A., & Kayser, I. (2012). Cannibalization or synergy? Consumers' channel selection in online–offline multichannel systems. *Journal of Retailing and Consumer Services, 19*(2), 186–194.

Kumar, V., Zhang, X., & Luo, A. (2014). Modeling customer opt-in and opt-out in a permission-based marketing context. *Journal of Marketing Research, 51*(4), 403–419.

Lee, M., Ha, T., Han, J., Rha, J. Y., & Kwon, T. T. (2015, June). Online footsteps to purchase: exploring consumer behaviors on online shopping sites. *Proceedings of the ACM Web Science Conference*, 1–10.

Leung, S. F., & Yu, S. (1996). On the choice between sample selection and two-part models. *Journal of Econometrics, 72*(1–2), 197–229.

Levin, N., & Zahavi, J. (2001). Predictive modeling using segmentation. *Journal of Interactive Marketing, 15*(2), 2–22.

Li, S., Sun, B., & Montgomery, A. L. (2011). Cross-selling the right product to the right customer at the right time. *Journal of Marketing Research, 48*(4), 683–700.

McDonald, J. F., & Moffitt, R. A. (1980). The uses of Tobit analysis. *The review of economics and statistics*, 318–321.

Meire, M., Hewett, K., Ballings, M., Kumar, V., & Van den Poel, D. (2019). The role of marketer-generated content in customer engagement marketing. *Journal of Marketing, 83*(6), 21–42.

Milne, G. R., & Boza, M.-E. (1999). Trust and concern in consumers' perceptions of marketing information management practices. *Journal of Interactive Marketing, 13*(1), 5–24.

Montgomery, A., & Smith, M. (2009). Prospects for personalization on the internet. *Journal of Interactive Marketing, 130–37*.

Moriarty, R. T., & Swartz, G. S. (1989). *Automation to Boost Sales and Marketing*. Harvard Business Review, Reprint Service.

Moriguchi, T., Xiong, G., & Luo, X. (2016). *Retargeting Ads for Shopping Cart Recovery: Evidence from Online Field Experiments*. Available at SSRN https://ssrn.com/abstract=2847631. or , https://doi.org/10.2139/ssrn.2847631.

Murray, K. B., & Häubl, G. (2009). Personalization without interrogation: towards more effective interactions between consumers and feature-based recommendation agents. *Journal of Interactive Marketing, 23*(2), 138–146.

Neslin, S., Grewal, D., Leghorn, R., Shankar, V., Teerling, M. L., & Thomas, J. S., et al (2006). Challenges and opportunities in multichannel management. *Journal of Service Research, 9*(2), 95–112.

Neslin, S., Powell, S. G., & Schneider Stone, L. (1995). The effects of retailer and consumer response on optimal manufacturer advertising and trade promotion strategies. *Management Science, 41*(5), 749–766.

Nowak, G. J., & Phelps, J. (1995). Direct marketing and the use of individual-level consumer information: determining how and when 'privacy' matters. *Journal of Direct Marketing, 9*(3), 46–60.

Park, Y.-H., & Fader, P. S. (2004). Modeling browsing behavior at multiple websites. *Marketing Science, 23*(3), 280–303.

Postma, O. J., & Brokke, M. (2002). Personalization in practice: the proven effects of personalization. *Journal of Database Marketing, 9*(2), 137–142.

Puhani, P. (2000). The Heckman correction for sample selection and its critique. *Journal of Economic Surveys, 14*(1), 53–68.

Reibstein, D. J. (2002). What attracts customers to online stores, and what keeps them coming back? *Journal of the Academy of Marketing Science, 30*(4), 465.

Rossi, P. E., McCulloch, R. E., & Allenby, G. M. (1996). The value of purchase history data in target marketing. *Marketing Science, 15*(4), 321–340.

Saha, A., Capps Jr., O., & Byrne, P. J. (1997). Calculating marginal effects in models for zero expenditures in household budgets using a Heckman-type correction. *Applied Economics, 29*(10), 1311–1316.

Sahni, N. S., Narayanan, S., & Kalyanam, K. (2019). An experimental investigation of the effects of retargeted advertising: The role of frequency and timing. *Journal of Marketing Research, 56*(3), 401–418.

Sahni, N. S., Wheeler, S. C., & Chintagunta, P. (2018). Personalization in email marketing: The role of noninformative advertising content. *Marketing Science, 37*(2), 236–258.

Schumann, D. W., Petty, R. E., & Scott Clemons, D. (1990). Predicting the effectiveness of different strategies of advertising variation: A test of the repetition-variation hypotheses. *Journal of Consumer Research, 192–202*.

Smith, M. D. (2002). The impact of shopbots on electronic markets. *Journal of the Academy of Marketing Science, 30*(4), 446.

Song, J. H., Kim, H. Y., Kim, S., Lee, S. W., & Lee, J. H. (2016). Effects of personalized e-mail messages on privacy risk: Moderating roles of control and intimacy. *Marketing Letters, 27*(1), 89–101.

Srivastava, V, & Kalro, A. D (2019). Enhancing the helpfulness of online consumer reviews: The role of latent (content) factors. *Journal of Interactive Marketing, 48*, 33–50.

Steinker, S., Hoberg, K., & Thonemann, U. W. (2017). The value of weather information for e-commerce operations. *Production and Operations Management, 26*(10), 1854–1874.

Telang, R., Boatwright, P., & Mukhopadhyay, T. (2004). A mixture model for internet search-engine visits. *Journal of Marketing Research, 41*(2), 206–214.

Todri, V., Ghose, A., & Singh, P. V. (2020). Trade-offs in online advertising: Advertising effectiveness and annoyance dynamics across the purchase funnel. *Information Systems Research, 31*(1), 102–125.

Urban, G. L., Hauser, J. R., Liberali, G., Braun, M., & Sultan, F. (2009). Morph the web to build empathy, trust and sales. *MIT Sloan Management Review, 50*(4), 53–61.

Vafainia, S., Breugelmans, E., & Bijmolt, T. (2019). Calling customers to take action: the impact of incentive and customer characteristics on direct mailing effectiveness. *Journal of Interactive Marketing, 45*, 62–80.

Vance, C. (2009). Marginal effects and significance testing with Heckman's sample selection model: a methodological note. *Applied Economics Letters, 16*(14), 1415–1419.

Verhoef, P. C., Neslin, S. A., & Vroomen, B. (2007). Multichannel customer management: Understanding the research-shopper phenomenon. *International Journal of Research in Marketing, 24*(2), 129–148.

Verhoef, R. V., McAlister, L., Malthouse, E., Krafft, M., & Ganesan, S. (2010). CRM in data-rich multichannel retailing environments: A review and future research directions. *Journal of Interactive Marketing, 24*, 121–137.

Vesanen, J. (2007). What is personalization? A conceptual framework. *European Journal of Marketing, 41*(5/6), 409–418.

Wattal, S., Telang, R., Mukhopadhyay, T., & Boatwright, P. (2012). What's in a "name"? Impact of use of customer information in e-mail advertisements. *Information Systems Researchs, 23*(3-part-1), 679–697.

White, T. B., et al (2008). Getting too personal: reactance to highly personalized email solicitations. *Marketing Letters, 19*(1), 39–50.

Wu, J., Li, K. J., & Liu, J. S. (2018). Bayesian inference for assessing effects of email marketing campaigns. *Journal of Business & Economic Statistics, 36*(2), 253–266.

Zhang, J., & Krishnamurthi, L. (2004). Customizing promotions in online stores. *Marketing Science, 23*(4), 561–578.

Zhang, & Wedel, M. (2009). The effectiveness of customized promotions in online and offline stores. *Journal of Marketing Research, 46*(2), 190–206.

Zhang, X., Kumar, V., & Cosguner, K. (2017). Dynamically managing a profitable email marketing program. *Journal of Marketing Research, 54*(6), 851–866.

Zhelonkin, M., Genton, M. G., & Ronchetti, E. (2016). Robust inference in sample selection models. *Journal of the Royal Statistical Society: Series B: Statistical Methodology*, 805–827.