The Impact of Subscription Programs on Customer Purchases

Raghuram Iyengar, Young-Hoon Park, and Qi Yu*

December 2018

* Raghuram Iyengar is Miers-Busch W’1885 Professor and Professor of Marketing at the Department of Marketing at the Wharton School, University of Pennsylvania, Philadelphia, PA 19104; email: riyengar@wharton.upenn.edu. Young-Hoon Park is Sung-Whan Suh Professor of Management and Professor of Marketing at the Samuel Curtis Johnson Graduate School of Management, Cornell University, Ithaca, NY 14853; email: yp34@cornell.edu. Qi Yu is a PhD student in Marketing at the Wharton School, University of Pennsylvania, Philadelphia, PA 19104; email: qiyu@wharton.upenn.edu. All authors contributed equally and are listed in alphabetical order. We thank several managers at the participating company for their efforts in making this research possible. We thank Ryan Dew, Pinar Yildirim, Christophe Van den Bulte and seminar participants at 2017 and 2018 Marketing Science Conference for their comments.
The Impact of Subscription Programs on Customer Purchases

Abstract

Subscription programs are increasingly popular among a wide variety of retailers including Amazon (Prime), Barnes & Noble (B&N Membership), and Sephora (Flash). These types of programs give members access to a set of exclusive benefits for a fixed fee upfront. In this paper, we document the causal effect of customers' adoption of a subscription program on their subsequent purchases. Using a unique data set of a subscription program launched by a firm in the fast-moving consumer goods industry, we find subscription leads to a very large increase in customer purchases, with significant variation across customers. The effect is economically significant and persistent. Members are more engaged with the company as they purchase more frequently and a greater variety of products. We provide evidence that customers experience a sunk cost fallacy and switch their purchases away from competitors to the focal company in product categories with lower switching costs. We discuss the implications for subscription businesses and customer retention.

Keywords: Subscription business, Causal inference, Machine learning, Causal forest, Retailing, E-commerce, Switching cost
1. **Introduction**

The ability to retain and develop customers is critical to the success of a business. Subscription programs, which are designed to keep customers engaged by giving access to exclusive benefits for a fee upfront, have become increasingly popular among retailers.\(^1\) For example, Amazon Prime offers members unlimited free two-day shipping, audio and video content, as well as member-exclusive discounts for an upfront payment of $119 per year. Many other retailers, such as Barnes & Noble, Sephora, and Alibaba have programs as well, and benefits range from unlimited free shipping to member-only discounts and additional loyalty points.

As the relevance and popularity of subscription programs has grown considerably in recent years, it is of managerial interest to quantify the effect of such programs on customer behavior and investigate underlying drivers for their apparent success. Customers may change their purchase behavior after becoming a member of a subscription program due to its economic benefits. Members may also respond to the program due to other psychological drivers. For example, the program could lead consumers to form habitual consumption, to feel enhanced status, and to generate positive affect for the company. Understanding the economic benefits that members deem important can assist managers to allocate resources accordingly. It is also informative to distinguish customer response between economic benefits and other psychological underpinnings. If additional sales are generated by reducing (effective) prices, the program might not be sustainable and negatively affect a firm’s profit in the long term (e.g., Raghurib 2004).\(^2\)

---

\(^1\) In this research, we distinguish subscription programs from stand-alone subscription services (e.g., Stitch Fix, Birchbox, Blue Apron) which provide subscribers new items or personalized experiences periodically. We focus on a setting where a subscription is initiated by an existing non-contractual business and provides members exclusive benefits beyond those available to regular customers (i.e., non-members). See Section 2 for a detailed discussion on different types of subscription programs.

\(^2\) For instance, Movie Pass, a subscription service that aimed to provide its members one free movie each day for the price of $9.95 per month, managed to attract more than 2 million subscribers with this low price but presumably failed to build a deeper relationship with the customers. As a result, the company reported a loss of $150.8 million in 2017, and was recently forced to alter their pricing strategies to keep the company afloat.
There are two major challenges when establishing the causal effect of a subscription program. First, it is likely that customers self-select into the membership. Consider Amazon’s Prime membership as an example. Anecdotal evidence suggests that Prime members spend $1300 per year, which is almost double the average non-member’s annual spending of $700.³ However, this naïve comparison likely over-estimates the effect of the program as customers who expect to make more purchases in the future are more likely to subscribe to Prime. Second, it is challenging to identify underlying drivers at work as subscription programs tend to offer multiple benefits and members may vary in how much they use these benefits. Without detailed information on how members benefit from each part of the program, it will be non-trivial to attribute the impact of the program to specific features, not to mention other behavioral explanations.

The purpose of this paper is to take a first step towards assessing the causal impact of customers adopting a subscription program on their subsequent purchase behavior. Specifically, we seek to examine the following questions: Is the subscription program effective in increasing customer purchases?, How does the effect vary over time and by customers?, What are underlying drivers? We address these questions in close collaboration with a company in the fast-moving consumer goods industry that launched a subscription program on its online website in 2015. For an upfront annual fee of $50, the benefits of the program include unlimited free shipping, and access to member-exclusive offers. Our data contains individual-level transactions before and after the launch of the program, as well as individual-level usage of various components in the program, thus presenting an appealing context for the study.

As noted earlier, a key concern while estimating the impact of a subscription program on purchase behavior arises from the lack of random assignment. We adopt the causal forest framework to address the issue of self-selection (Wager and Athey 2018). Specifically, we apply a random classification forest to the change in the variables of interest (e.g., purchase amount) pre- and post-subscription, match members of the program to non-members with similar characteristics, and compare the changes in purchase behavior within matched pairs. In essence, we combine a difference-in-difference estimation together with the causal forest framework to estimate the treatment effect, i.e., the impact of subscription program on customer purchases. The causal forest is robust to model mis-specification, non-parametric, and computationally efficient. It also provides individual-level estimates of the treatment effect, thus allowing us to get a richer understanding of the treatment effect and its possible underlying drivers.

We find that on average, members increase their purchases by $30 per month, a substantial increase of 150%, after joining the program. The effect on customer purchases is economically significant in the short run (an increase of $46 in the first month after joining) and persistent in the long run (an increase of $28 per month after three months upon joining). Members are more engaged with the company as they purchase more frequently and a greater variety of products. There is also substantial variation across customers on the impact of the program. For instance, customers with a lower purchase frequency in the pre-subscription period increase their purchases more so after joining the program. Similarly, customers who are willing to try new products and make repeat purchases of known products increase their purchases more so with the program. Our main results are robust to different subsamples of the treatment group and other definitions of matching samples.
We investigate potential mechanisms at work to explain our results. We find that cross-channel substitution and economic benefits of the subscription (e.g., unlimited free shipping, member-exclusive offers) cannot explain the large increase in customer purchases. Neither do common psychological underpinnings (e.g., habit, status, affect) of other types of membership programs (e.g., loyalty programs) rationalize our empirical findings. Rather, a unique feature of the subscription program helps to boost sales: as customers pay a fixed fee upfront in exchange for benefits in future, they can suffer from sunk cost fallacy in which they increase their purchases to justify their subscription decisions, even though the subscription fee is sunk (Thaler 1980). We provide evidence supporting this mechanism by demonstrating that the boost in sales is larger in product categories with lower switching costs.

Our paper is related to a few streams of research. First, we add to the literature on membership programs. Firms across a wide array of industries have long been using loyalty programs to reward repeat purchases, and there is extensive research on these types of programs. Several studies have found that loyalty programs can increase customers’ share of wallet and lifetime value (e.g., Lal and Bell 2003, Lewis 2004, Taylor and Neslin 2005, Liu 2007). Others found no or weak evidence that loyalty programs are effective (e.g., Hartmann and Viard 2008, Meyer-Waarden and Benavent 2009). Several researchers have documented that loyalty programs can lead to the development of habitual consumption (e.g., Wood and Neal 2009), enhance members’ perception of status (e.g., Drèze and Nunes 2009), and induce positive affect (Leenheer et al. 2007).

In this research, we study an emerging type of membership program, subscription programs, whose customers pay a periodically recurring fee for access to products or services. Recently, McCarthy, Fader, and Hardie (2017) develop a framework for valuing subscription-
based firms, and Datta, Knox, and Bronnenberg (2018) study how the adoption of music streaming affects listening behavior. Our paper contributes to this nascent and growing literature on two important dimensions. Substantively, while empirical evidence on customer response to subscription programs is still limited, our paper takes the first step by empirically examining the causal impact of a subscription program on customer purchases. Theoretically, we find that subscription programs differ from loyalty programs in terms of how they work. Because customers pay a subscription fee upfront, they desire to increase their purchases to take advantage of the program benefits in order to justify their subscription decisions.

Broadly, our results indicate that customers behave in a boundedly rational manner, which adds to the empirical evidence for such behavior found in the lab and in other field settings. The sunk cost fallacy, the general tendency for people to continue consuming or pursuing an option given their initial investment, has implications in a variety of contexts, and extensive evidence has been found in the lab (Thaler 1980). There are relatively few studies finding evidence for the sunk cost fallacy in the field. For example, Arkes and Blumer (1985) conduct a field experiment and find the price of theater tickets is positively correlated with the attendance rate. Recently, Ho, Png, and Reza (2017) find evidence for the sunk cost in the Singapore automobile markets where there is heterogeneity among consumers with regard to the payment for obtaining a government license to purchase a car. And the driving time is positively correlated with the price paid. We extend this stream of literature by showing that subscription programs also induce the sunk cost fallacy and contribute to our understanding of consumer behavior in the field using observational data.

The remainder of the paper is organized as follows. Section 2 gives an overview of subscription business. Sections 3 and 4 discuss the data and methodology. Section 5 presents the
results and discuss possible explanations for the effects. Section 6 presents several robustness checks. We conclude with directions for future search in Section 7.

2. Subscription Business

A subscription-based business is one where a customer pays a fee to have access to products or services. Rather than selling products individually, a subscription offers periodic (e.g., monthly, yearly) use or access to products or services. Thus, a one-time purchase of a subscription can lead to recurring sales, and a predictable stream of revenues from subscribers (Baxter 2015). Pioneered by the likes of newspapers and magazines, more products and services are being offered through subscriptions than ever before. For instance, business-to-consumer subscription businesses attracted more than 11 million U.S. subscribers in 2017, and the industry as a whole has been growing at a staggering rate of 200% annually since 2011.

Despite sharing a common feature of offering the use of or access to products or services for a fee, subscription-based businesses appear in many different formats. Existing business-to-consumer subscription business in retail can be broadly categorized into three types: replenishment, curation, and access.4

Replenishment subscriptions allow consumers to automate the purchase of commodity items, such as razors, diapers, and vitamins. With automatic replenishment, customers no longer need to take the time and effort to repeat the order themselves. Customers benefit from this type of subscription because it allows them to save time and money on each transaction. Examples

4 There are over 2,000 consumer-focused subscription businesses capitalizing upon customers’ diverse tastes in a wide range of categories. There are also hundreds of companies with more unorthodox products catering to the “long tail” of consumer tastes, including Harry Potter toys and survivalist products. As subscription-based businesses continue to grow, we recognize that the three categories described may not be sufficient to capture the diversity. Please see the article available at https://www.mckinsey.com/industries/high-tech/our-insights/thinking-inside-the-subscription-box-new-research-on-ecommerce-consumers (last accessed December 18, 2018).
include Dollar Shave Club, Gillette on Demand, and Rituals. Curation subscriptions seek to
delight by providing new items or personalized experiences in such categories as apparels,
beauty, and food. Examples include Stitch Fix, Birchbox, and Blue Apron.

In this paper, we focus on the third type of subscriptions, access subscriptions that allow
consumers to gain exclusive access or member-exclusive benefits. Access subscriptions have
attracted substantial interest among more established retailers, as compared to the first two types
of subscriptions that are mostly launched by start-ups fueled by venture capital investments.
Examples include Amazon (Prime), Barnes & Noble (B&N Membership), Sephora (Flash),
Alibaba (88VIP), etc. These subscriptions differ from the other two described above in that a
subscription is initiated by an existing non-contractual business and provides members exclusive
benefits beyond those available to regular customers (i.e., non-members). Thus, they have a very
wide appeal and could be adopted by nearly all business-to-consumer firms. While there are
many variations of such programs in practice, benefits offered to members typically fall into two
broad categories: unlimited use of a service (e.g., unlimited free shipping) and access to member-
exclusive offers. Amazon (Prime) and Barnes & Noble (B&N Membership), for example, offer
both types of benefits to their members while Sephora’s Flash offers only unlimited free
shipping. Firms also vary by the type of member-exclusive offers. Barnes & Noble provides
members with exclusive offers only for purchasing products, while Amazon and Alibaba
(88VIP) offer exclusive digital content to their members as well as member-only benefits related
with product purchases. For the remainder of the paper, we refer to “access subscriptions” as
subscriptions for brevity and use the terms “subscription program” and program interchangeably.

While there is anecdotal evidence suggesting the commercial success of subscriptions,
limited research has been conducted to evaluate whether they indeed lead to incremental
revenues from members as compared to non-members who can purchase at a firm without being subscribers, and to explore potential drivers at work. Our research aims to fill in this gap.

3. Empirical Context and Data

We obtained the data for our empirical analysis from a retailer in Asia. The retailer sells a wide range of beauty products (e.g., skin care, make-up) and has both brick-and-mortar and online presence. As the beauty industry is increasingly being characterized by diminishing product differentiation and heightened competitive pressure, retailers are exploring new ways to increase customer retention and value. As an avenue to achieve these goals, the retailer that we collaborate with launched a subscription program only on its online website in December 2015. The launch of the program and its benefits were communicated to online customers through mass emails and on their website. The program provides a wide range of benefits, typical to access subscriptions, covering both unlimited use of a service and access to member-exclusive offers. Hence, it allows us to examine the impact of different program benefits on customer behavior.

Specifically, for a subscription fee of $50 per year, members have access to unlimited free shipping service with no minimum purchase requirements. Members also have access to several exclusive offers. Upon signing up for the program, customers are provided with a $50 gift card that can be used for purchases online without any restrictions. Members are also offered a $3 gift card per month for online purchases during the month. Free samples are offered to members monthly with a purchase online. Finally, several products are occasionally coupled with member-exclusive discounts.

Our data includes the census of all the 21,264 customers who joined the program between December 2015 and February 2017. The data consists of two parts: transaction data and program usage data. The transaction data contains detailed individual-level information on each transaction, i.e., when a customer purchased a product (e.g., a moisturizer), and how much she paid for it. The program usage data contains rich individual-level information on how a member benefits from the program, e.g., amount spent with gift cards, free samples received, etc. This data allows us to disentangle the effects of different components of the program on customer behavior. Moreover, it also helps make inferences for underlying drivers at work. Our data also contains individual characteristics of customers, e.g., age, gender and home address, which further controls for customer heterogeneity.

Figure 1 shows the number of customers joining the program each month. On average, 720 customers joined the program per month. The monthly number of members has moderate level of variation, ranging from a minimum of 342 to a maximum of 1062, and with a standard deviation of 210.

Insert Figure 1 about here

Using transaction data, we define a set of metrics associated with customer purchases. As the program is only at the online channel, unless specified otherwise, these measures are based on online purchases where the program could have a direct effect, and are constructed at the customer-month level, which is the unit of analysis. As our primary interest is to assess how effective the program is in lifting sales, our primary measure is the purchase amount spent by a customer per month. In addition, we consider the number of purchases made and basket size conditional on purchase, because the increase in monthly purchase amount could be due to the increase in purchase frequency and/or basket size.
We also try to characterize the variety in purchase behavior with a few metrics. We classify a product (and its product category) a customer purchases as a new versus known product (category) on the basis of whether or not she had purchased it in the pre-subscription period (to be described in Section 4.2). The first set of metrics relates to the variety at the product level: purchase amount spent for new vs. known products. The second set of metrics relates to the variety at the category level: purchase amount spent for new vs. known categories. As a proxy for engagement to the firm, these measures are useful to investigate how the variety of new (vs. known) products and categories change when customers join the program.

4. Method

4.1 Identification Strategy

Our objective is to identify the causal effect of customers adopting the program on their purchase behavior. If we were able to conduct controlled field experiments in which customers are randomly assigned to the program, it would be relatively straightforward to determine the causal impact of the program. In this case, the outcomes of the assigned non-members would serve as the counterfactual outcomes of the treated group. However, such randomized experiments for assigning membership were not feasible in the course of the commercial operation of the company, implying that we have to approach this causal inference problem using observational data that lacks a randomized assignment.

Skin care, make-up, hair care, bath and body care, fragrance, etc. are the main product categories of beauty products. Based on conversations with the data provider, we decided to have five product categories for our empirical analysis in order to correspond to the way in which the firm monitors key metrics regarding customer purchases. They include skincare, make-up, hair care, bath and body care, and others in which we aggregated and grouped fragrance and the rest of the categories (e.g., tools and brushes).
Observational data, however, are challenging to use for the purpose because customers who expect higher purchases in the future are more likely to join the program. Without a valid control group, we cannot conclude that the program has a causal impact on customer behavior. Therefore, we rely on a quasi-experimental design to identify the causal effect of subscription on customer purchases. The idea is to estimate the treatment effect by considering comparable customers in the covariate space as if these customers came from a randomized experiment. The validity of this identification strategy then hinges on the assumption that conditional on the observed covariates, there is no selection on the unobservables.

4.2 Treated and Control Groups

We focus on a cohort of customers who join the program around the same time as it gives a well-defined pre- and post-subscription periods for the analysis. Focusing on a cohort is especially helpful as our goal is to study the causal effect of the program by comparing members with non-members who have similar characteristics. In order to mitigate the concerns for selection and unmeasured confounders, we deliberately exclude early adopters as they may systematically differ from other customers (e.g., Rogers 2003). Our main results consider the cohort of 721 members who joined the program in April 2016, four months after the launch of the program. For the purpose of comparison, we obtain a sample of 13,768 customers who have yet to subscribe to the program as of March 2017. We note that our findings are not specific to the cohort of members that we choose to focus on. As part of robustness checks, we estimate the effect of the program among members who joined the program in other months, and report in Section 6 that our main results are qualitatively similar and the effects are robust.

Before we establish the causal effect of customers adopting the program on customer purchases, as a first step, we examine the purchase amount for members and non-members over
24 months of the data period, April 2015 to March 2017. Of these, the first 12 months (i.e., April 2015-March 2016) are prior to the subscription. As shown in Figure 2, purchase behavior differs considerably between members and non-members. On average, members spend $43.16 per month post-subscription while non-members spend only $3.93. This cross-sectional comparison might suggest that the value of the program is $39.23 per month ($p < 0.001). However, it is likely that customers with different pre-subscription characteristics self-select into the program.

**Insert Figure 2 about here**

To assess whether non-members have the same characteristics as members prior to joining the program, we compare their purchases during the pre-subscription 12 months and individual characteristics. Table 1 shows that members and non-members differ significantly on their purchases and demographics. Members, on average, spend more per month than non-members in the pre-subscription period ($11.99 versus $6.37, $p < 0.001). This result is consistent with the intuition that customers who have higher purchases are more likely to join the program because they can benefit more from the program. Members are also older than non-members (35.94 versus 32.84, $p < 0.001). Clearly, estimating the effect of the program by merely comparing customer purchases between the two groups will be biased.

**Insert Table 1 about here**

One improvement on the naïve cross-sectional comparison is to construct a difference-in-difference estimator by comparing the change in customer purchases over time for the members with that for the non-members. In this way, any time-invariant heterogeneity could be removed. Such comparison leads to an estimate of the treatment effect of $32.90 ($p < 0.001). However, note again from Figure 2 and Table 1 that the difference in purchase amount between the two groups widens over time, suggesting that compared to an average non-member, the monthly
spend by members is increasing over time even before their subscriptions. Thus, the parallel time
trend assumption does not hold and the validity of a difference-in-difference estimator is
questionable (e.g., Bertrand, Duflo, and Mullainathan 2004).

4.3 Causal Forest

To achieve the comparability of customers between the treated and control groups, we employ
the causal forest method (Wager and Athey 2018). The causal forest method uses a random
classification forest to match the treated and control groups that share similar covariates. The
treatment effect is then estimated by comparing the outcome (e.g., purchase amount) of the
treatment unit and the “matched” control unit. As compared to other commonly used methods for
causal inferences using observational data, e.g., nearest neighbor matching (e.g., Imbens and
Rubin 2015), the adaptive nature in trees can substantially increase the power of accurate
clustering with a large space of covariates. The tree structure and the ensemble of many trees
also naturally account for complex interplay among covariates. Compared to propensity score
matching (e.g., Hirano, Imbens, and Ridder 2003) and synthetic control method (e.g., Abadie,
Diamond, and Hainmueller 2010), the causal forest method is robust to model mis-specification,
non-parametric, and computationally efficient. Another advantage of the causal forest is that it
achieves desired consistency and asymptotic normality when estimating the heterogeneous
treatment effects at the individual level. Any variation in the treatment effect across individuals
can sharpen our understanding of underlying drivers for the success of the program.

To enhance the robustness of our inference, instead of using the outcome itself, we apply
the causal forest to the change in the outcomes before and after a customer joins the program. In
essence, we combine a difference-in-difference estimation together with the causal forest
framework. This combination of estimation methods has the advantage of eliminating the time-
invariant unobserved heterogeneity that the observables fail to capture. It gives a better estimate for the treatment effect with less bias (e.g., Smith and Todd 2005).

Let \( \{ Y_i, X_i, W_i \} \) denote the data that the researcher observes, where \( Y_i \) denotes the outcome of interest for customer \( i \), \( X_i \) is the set of covariates, and \( W_i \) is the treatment indicator. Let \( Y_i^{(1)} \) and \( Y_i^{(0)} \) denote the potential outcomes that customer \( i \) would have with and without the treatment, respectively. The causal inference problem in observational studies can be expressed as estimating the conditional average treatment effect\(^7\):

\[
\tau(x) = E \left[ Y_i^{(1)} - Y_i^{(0)} \mid X_i = x \right]. \tag{1}
\]

The causal forest algorithm grows a classification tree by iteratively partitioning the data into subgroups (leaves). In each iteration, the algorithm chooses the covariate and the best cutoff by optimizing the Gini criterion (Breiman 2017) and continues to partition the data until there is at least one treatment and one control in each leaf. As a single tree is likely to overfit the data, an ensemble of \( B \) trees is generated. The \( b \)th tree is constructed using a random subsample without replacement containing \( n_b \) observations from a total of \( N \) observations in the data. For each tree, a random set of 1/3 of the covariates can be potentially used to form the splits.\(^8\)

Given the forest, we estimate the treatment effect by comparing the difference in the outcomes of the treatment unit and the “matched” control unit. We assume that unconfoundedness holds: \( W_i \perp \left\{ \left( Y_{iT_1}^{(1)} - Y_{iT_0}^{(1)} \right), \left( Y_{iT_1}^{(0)} - Y_{iT_0}^{(0)} \right) \right\} \mid X_i \), where \( T_0 \) and \( T_1 \) are the pre- and post-subscription period, respectively. The treatment effect \( \hat{\tau}_b(x) \) from the \( b \)th tree is given

\(^7\) We use the term “conditional average treatment effect” and “treatment effect” interchangeably.
\(^8\) The number of trees, the minimum number of treatment and control observations in each leaf, the subsample size, and the size of the set of covariates used to build each tree are the hyper-parameters of a causal forest that are typically chosen by the researcher. In the absence of a formal guidance of the choices of the hyper-parameters, we use a large number of trees to generate robust estimates and use a small number of observations in each leaf to encourage the discovery of heterogeneity. The subsample size and the size of the set of covariates used to build each tree are moderate and chosen to balance the bias-variance tradeoffs (Davis and Heller 2017).
by taking the difference between the outcomes of the treated and control groups within the same leaf $L$:

$$\hat{b}(x) = \frac{1}{|\{i: W_i = 1, X_i \in L\}|} \sum_{i: W_i = 1, X_i \in L} \left( \frac{1}{T_1} \sum_{t \in T_1} Y_{it} - \frac{1}{T_0} \sum_{t \in T_0} Y_{it} \right) - $$

$$- \frac{1}{|\{i: W_i = 0, X_i \in L\}|} \sum_{i: W_i = 0, X_i \in L} \left( \frac{1}{T_1} \sum_{t \in T_1} Y_{it} - \frac{1}{T_0} \sum_{t \in T_0} Y_{it} \right),$$

where $Y_{it}$ is the outcome variable for customer $i'$ in month $t$ and $\left( \frac{1}{T_1} \sum_{t \in T_1} Y_{it} - \frac{1}{T_0} \sum_{t \in T_0} Y_{it} \right)$ is the difference in purchases for customer $i$ between the post- and pre-subscription periods.

To reduce bias and ensure correct inferences, Wager and Athey (2018) recommend an “honest” approach, which uses different subsamples for growing trees and for estimating treatment effects. For a tree $b$, another subsample of size $n'_b$, which is disjoint from the subsample $n_b$ used to place splits, is used in order to estimate the treatment effect. Growing the trees and estimating the treatment effects using different subsamples reduces the correlation between the tree structure and the estimated treatment effect, and the consistency and centered asymptotic normality of the estimates are guaranteed.

With an estimate $\hat{b}(x)$ from the $b$th tree, we average over the estimates across $B$ trees and have the final estimate of the conditional average treatment effect:

$$\hat{f}(x) = \frac{1}{B} \sum_b \hat{b}(x).$$

To measure the randomness in $\hat{f}(x)$ due to the training sample, we estimate the variance of the causal forest as follows:

$$\hat{V}(x) = \frac{N-1}{N} \left( \frac{N}{N-(n_b+n'_b)} \right)^2 \sum_{i=1}^{N} \text{cov}(\hat{b}(x), I_{ib})^2.$$

---

$^9$ We further require each leaf of the tree to contain at least one treatment and one control from the latter sample to enable the estimation of the treatment effects.
where \( \frac{N-1}{N} \left( \frac{N}{N-(n_b+n_b^i)} \right)^2 \) is a finite-sample correction for forests grown by subsampling without replacement and \( I_{ib} \in \{0,1\} \) indicates whether observation \( i \) is used for the \( b \)th tree. The covariance is taken with respect to all \( B \) trees in this forest.

We next describe implementation details. For the outcome variables, we use 24 months of transaction data (April 2015 to March 2017) associated with the program. Of these, 12 months are prior to the subscription. In our estimation, we exclude customer purchases in the month of April 2016 to avoid any potential simultaneity bias with the adoption itself.\(^{10}\) The covariate vector \( X_t \) contains three sets of variables. The first set of covariates relates to the customer-firm relationship which would be associated with the adoption of a service (e.g., Bolton, Lemon, and Verhoef 2004, Prins and Verhoef 2007): tenure, breadth, and depth. We calculate tenure based on elapsed time since having an account on the website. We measure breadth by the number of unique categories purchased and depth by the number of transactions made. We also include the average basket size. In addition, we include monthly purchase amount per category during the 12-month pre-subscription period, instead of the total amount across product categories, because it could help find clusters of customers with similar purchase sequences and patterns across different product categories. We also include the variance of monthly purchase amount that could be related to customer response for unlimited free shipping service.

The second set of covariates relates to psychographic measures that reflect personality traits, preferences or interests, values and attitudes (e.g, Baumgartner 2002). Because psychographics is not directly observable, their measure is less precise and more nuanced. But, it could capture underlying customer psychographics that impact subscription and contribute to the

---

\(^{10}\) The estimate we report provides a conservative estimate of the effect on customer purchases. Detailed results which include April 2016 are available from the authors upon request.
identification of customer clusters (e.g., Müllensiefen, Hennig, and Howells 2018). Given that
we use observational data with no surveys sent to customers, we explore a few measures
associated with customer purchases by summarizing certain aspects of purchase behavior (e.g.,
Baumgartner 2002): exploratory (purchases of new products), repetitive (repeat purchases of a
product), and promotional (discounts received for purchases).\footnote{We conduct our analysis with different variable operationalizations and find that the results are qualitatively similar to the results reported.} The third set of covariates relates
to socio-demographics of customers. We include age and gender. We also include the
coordinates of a customer’s home address, because it can help control for other unobserved
socio-demographics that affect subscription, e.g., education, income, household size, life style,
etc. Table 2 summarizes the covariate vector $X_t$ and describes how the variables are
operationalized. Altogether, we use 73 covariates to build the trees for the causal forest.

\underline{Insert Table 2 about here}

The forest performs well in classifying members and non-members.\footnote{We fix $n_b$ (and $n'_b$) to be $0.1N$ and $B$ to be 2,000 in the estimation.} The Area Under
the Curve (AUC) of the forest is 0.91. Notably, consistent with prior literature, the customer-firm
relationship and psychographics variables constructed from prior purchases have the most
predictive power, while socio-demographics have relatively low predictive power in explaining
the adoption of subscription. It also shows the importance of conditioning on a rich set of
covariates to achieve the comparability between the treatment and the matched control units.

5. Findings
We begin by discussing the main findings on the treatment effect of the program on customer purchases. We then discuss the heterogeneity in the effects across customers and offer some explanations for the drivers.

5.1 Treatment Effects

In Table 3, we report the treatment effect and 95% interval of the program on purchase amount and other metrics associated with customer purchases. On average, purchase amount per month among members increases by $29.82.13 The magnitude of the increase in monthly expenditure is quite striking as average purchase amount per month is only about $12 before the subscription.

We investigate how the treatment effect on purchase amount varies over time. The program could create an initial excitement among members, leading to their increased purchases. If this is the primary reason for their behavioral change, the effect might fade away over time and the program would have limited impact on the firm’s long-term revenues. To examine the temporal effect of the program on customer purchases, we decompose the treatment effect into the short-, medium- and longer-run effects. We define short-run as one month, medium-run as two to three months, and longer-run as more than three months post subscription. We estimate how purchase amount in these three different time windows changes relative to the 12-month pre-subscription period by applying the causal forest algorithm. Our results show that, as expected, the effect is the largest in the short term. Purchase amount increases by $46.20 in the short term. The effect decreases marginally in the medium term ($27.34). The effect for the longer term is significant at $28.33 per month. In sum, the effect of the program on purchase amount is economically significant not only in the short run but also persistent in the long run.

13 Recall that the cross-sectional and difference-in-difference estimates are $39.23 and $32.90, respectively. Comparing the treatment effect obtained from the causal forest, one could over-estimate the effect of the program on monthly spend by about $10 (33%) with cross-sectional estimate and by about $3 (10%) with difference-in-difference estimate.
The increase in customer purchases could be driven by the increase in purchase frequency and/or basket size. These two metrics, although typically positively correlated, could have different implications for the firm. Similar to the estimation with purchase amount, we estimate the causal forest model using the differences in each of these two variables as an outcome measure. Our results show that members make about one additional purchase per month (1.01) post subscription. And there is a small decrease in the basket size (-4.96), likely due to unlimited free shipping. Taken together, our evidence supports that subscription is effective in lifting sales in the post-subscription period and does so by making members more frequent shoppers than appropriately matched non-members.

**Insert Table 3 about here**

We next examine the effect of subscription on the variety measures. Recall that we classify products (categories) a customer purchases to new vs. known products (categories) based on prior purchase behavior. We find a significant increase in all measures. At the category level, over 80% of the increase in purchase amount ($24.25 out of $29.82) comes from known categories, and the rest from new categories that a customer had never purchased in the pre-subscription period. At the product level, approximately 70% of the increase in purchase amount ($20.81 out of $29.82) comes from new products that a customer had never purchased. These results strongly point to an increase in the variety not only at the category level but also at the product level, leading to increased customer loyalty and share of wallet.

In summary, we conclude that there is indeed a causal impact of subscription program on customer purchases. The effect of the program on customer purchases is economically and managerially significant, and holds not only for the short run but also for the longer run. The program keeps customers more engaged both in terms of frequency and variety.
5.2 Heterogeneous Treatment Effects

Figure 3 shows the distribution of the treatment effect in purchase amount for members. We observe large variation in the changes in purchase amount across members, ranging from less than $10 to more than $100. About 25% of the members increased their purchases by $15 or less. In contrast, about 30% increased their purchases by $40 or more, and approximately 5% increased by $80 or more. This result also illustrates the benefits of employing the causal forest framework that the treatment effects are estimated at the individual level.

**Insert Figure 3 about here**

We next examine how individual characteristics moderate the impact of the program on customer purchases. Traditionally, researchers search for subgroups with different treatment effects in an ad hoc manner. The heterogeneity discovered in this way, usually in the form of a few of interactions, may be spurious (e.g., Cook, Gebski, and Keech 2004). As the causal forest estimates the treatment effects at the individual level with asymptotic standard errors, we can explore the heterogeneous treatment effects in a systematic manner. To this end, we relate the treatment effects to the observed customer characteristics by regressing individual-level treatment effects onto observed characteristics and estimate the following specification:

\[
\hat{\tau}_i = \beta_0 + \beta_1 X_i + \epsilon_i,
\]

where \(\hat{\tau}_i\) is the estimated treatment effect of the program on purchase amount for customer \(i\) and \(X_i\) is a vector of individual characteristics that include the variables described in Table 2 (i.e., customer-firm relationship, psychographics, and socio-demographics).\(^{14}\) To improve the interpretability of the results, we use monthly purchase amount across product categories, instead of the one at the category level, in the regression. As the dependent variable in the above

\(^{14}\) We exclude the coordinate of home address from the regression.
regression is estimated with error, we use generalized least squares to estimate the model to account for the variance of the dependent variables (Hanushek 1974).

Table 4 reports the results. Our results suggest that customers with lower frequency likely increase their purchases more after joining the program. Also, the program has a larger impact on newer customers. Interestingly, our results suggest that customers with higher frequency and longer relationship with the firm increase their purchases only moderately. Customers who try new products and make repeat purchases of products increase their purchases more with the program. These results can assist marketers in scoring customers and making targeting decisions across individuals for marketing activity associated with subscriptions.

Insert Table 4 about here

5.3 Possible Explanations

In this section, we explore potential drivers for the success of the online subscription program. As the firm we partnered with has both brick-and-mortar and online presence, we first investigate whether customers are simply moving their purchases from offline to online. We next utilize the program usage data and examine which benefit of the program, or a combination of them, contributes to increased purchases. Finally, drawing from the behavioral literature and also using individual-level estimates of the treatment effect, we investigate four possible behavioral underpinnings for the effect of the program on customer purchases.

5.3.1 Offline Purchases

As the subscription program we study is an online-only program, our previous analysis focuses on customer purchases on the website. For multi-channel firms, depending on the nature of their products and consumers, their online and offline channels could be either complements or substitutes (e.g., Forman, Ghose, and Goldfarb 2009, Wang and Goldfarb 2017). Thus, one
possible explanation for the changes in purchases online is that customers might simply switch their purchases from offline to online. If this were the case, the program might not have any effect on customer purchases in total and the program will not increase revenues for the firm.

We examine this explanation by evaluating the effect of the program on customer purchases offline. To further control for the fact that members and non-members might be different in their offline purchases, we redo the matching of members and non-members with pre-subscription purchase behavior across the two channels, along with all the other observed characteristics applied in our main analysis. We find no significant changes in customer purchases at the offline channel (-$2.69, 95% interval [-14.36, 6.24]). This result suggests that in our context, the online and offline channels are only weak substitutes. Therefore, the program is effective in lifting overall revenues for the firm.

5.3.2 Economic Benefits
Members receive two types of economic benefits: unlimited free shipping and member-exclusive offers. To examine the role each part of the program played in generating incremental sales, we make use of program usage data. This data contains rich information on gift cards used, free samples offered, and shipping fees waived. We also use transaction data which includes the prices listed and paid at the customer-product level, and calculate any discounts that members received. All of these data allow us to examine how members benefit economically from the program.

As expected, most members (91%) use up their $50 gift cards that were offered upon joining the program.15 Interestingly, only about 20% of monthly $3 gift cards are utilized,

---

15 Members also lose fungibility of the $50 gift card. However, we note that the loss of fungibility does not affect customer purchases for two reasons. First, most members spend more than $50 per year. Second, about half of the members use up their $50 gift cards in the month they join the program. Since we find that the program has a persistent long-term effect, the loss of fungibility cannot explain the increase in customer purchases.
probably because members do not make purchases online every month. To explore how an additional $3 would affect customer purchases, we leverage prior findings that the elasticity of store coupons on purchase amount is less than 1 (e.g., Kumar and Swaminathan 2005). While redemption behaviors likely differ between online and offline contexts, the increase in customer purchases due to $3 gift cards would likely be less than $3, i.e., net benefit. Hence, monthly gift cards alone do not explain the increase in customer purchases.

Free samples are another avenue for the increase in customer purchases. Similar to monthly gift cards, only about 30% of free samples are redeemed. To explore the impact of free samples on customer purchases, we differentiate the purchases induced by free samples from others. We define a purchase to be induced by a free sample if a customer had received a free sample of the product in the past. Note that this measure is chosen to investigate an upper bound on the effect of free samples, as the customer can purchase the product without trying the sample. Nevertheless, we find that less than 1% of purchases are induced by free samples.

Member-exclusive discounts, although occasional, are another possible explanation for the increase in customer purchases. To investigate the impact of member-only discounts on purchases, ideally, one would compare customer purchases with and without these discounts. While transaction data contains information on the prices listed and paid by a customer when she buys a product, we do not observe the counterfactual price she would have faced had she not joined the program. To address this issue, we use the average discount rate non-members obtain as a proxy for the counterfactual level of discount rate a member would have encountered had she not joined the program. Our results show that on average, members were given 6% discounts more than non-members. If the increase in purchases is solely explained by member-exclusive
discounts, it would imply a price elasticity that is atypically large. Thus, member-exclusive
discounts are less likely to explain the increase in customer purchases.

Finally, the online website offers customers, regardless of the subscription, free shipping
on orders of $10 or more, and less than 0.1% of the orders fail to meet that threshold. Hence, the
impact of unlimited free shipping itself is quite limited.

The results discussed so far suggest that each of the economic benefits alone only
explains a small part of the increase in customer purchases. We now investigate whether all the
economic benefits of the program can collectively explain the increase in customer purchases. To
this end, we estimate the following model:

\[ \text{Purchase}_{it} = \beta_1 \text{Member}_{it} + \beta_2 \text{Price}_{it} + \beta_3 \text{Giftcard}_{it} + \alpha_i + \gamma_t + \epsilon_{it}, \]

where \( \text{Purchase}_{it} \) is the amount spent by customer \( i \) at month \( t \) that is not induced by free
samples. \( \text{Member}_{it} \) is a dummy variable that indicates whether or not customer \( i \) is a member of
the program at month \( t \). We include the average price customer \( i \) pays at time \( t \) (\( \text{Price}_{it} \)), which
controls for the fact that members and non-members encounter different prices when making
their purchase decisions, and the amount of gift cards a member redeems in a given month
(\( \text{Giftcard}_{it} \)). We also include individual- and time-fixed effects. The results from the regression
indicate a positive and significant \( \hat{\beta}_1 \) (24.28, \( p \)-value < 0.001). We conclude that even after
controlling for all the economic benefits of the program, a significant part of the impact of the
program on purchases is left unexplained.

5.3.3 Beyond Economic Benefits

In this section, we propose four possible explanations for the effect of the program on customer
purchases drawing from the behavioral literature. We also present additional empirical patterns
that may help explain underlying behavior among members. However, we caution that it is
challenging to conclusively pinpoint one explanation with observational data and there can even be multiple forces at work. Hence we conclude by pointing to the explanation that collectively rationalizes our empirical findings.

The first three explanations we propose are not specific to subscription programs. In fact, any other membership programs, including loyalty programs, can show similar effects (see Bolton, Lemon, and Verhoef 2004 for a review). First, the program can encourage customers to purchase after subscription, leading to habitual purchases in the long run (e.g., Wood and Neal 2009). Second, members may feel superior when they have access to several exclusive offers, and their enhanced status can encourage purchases (e.g., Drèze and Nunes 2009). Third, the program can generate positive affect towards the firm and lead to increased purchases (e.g., Leenheer et al. 2007). Finally, as customers pay a fixed fee upfront in exchange for benefits in future, they can suffer from sunk cost fallacy in which they increase their purchases to justify their subscription decisions, even though the subscription fee is sunk (Thaler 1980).

We investigate the first explanation (habit) by comparing the purchase behavior of members with that of non-members that have similar purchases not only before subscription but also shortly after subscription. The rationale is that by matching on post-subscription purchases, members and (matched) non-members should then be comparable in terms of consumption patterns till after subscription. Hence any difference in their future purchases must stem from reasons other than their consumption habit. Specifically, we build the causal forest using all the variables in the main analysis as well as purchase amounts in five categories (i.e., skincare, make-up, hair care, bath and body care, and others) till July 2016 (three months after members join the program). We then estimate the changes in monthly purchases after August 2016 relative to the pre-subscription period. Using this specification, we find that members increase their
purchases by $25.57 per month. This result suggests that the treatment effect is not likely driven by the state dependence of purchases.

Our approach to test the second explanation (status) is to re-examine the effect of the program over time. The value of status created by membership is associated with its distinctiveness (e.g., Grier and Deshpandé 2001). As more customers subscribe, the program would provide members with decreasing value. If members derive social status from the program, we would expect that the program should have a smaller impact on their purchases in the long run. While we do find that the short-term effect of the program is the largest, the medium- and longer-term effects of the program are quite indistinguishable. We interpret this temporal pattern of results as somewhat weak evidence for the status explanation.

The third explanation (positive affect) would predict that customers with low engagement with the firm would have a higher increase in purchases after they become members of the program. Further, as engaged customers already have favorable affect towards the firm, the effect of the program could be limited for them. This explanation is largely consistent with our findings that the program has a larger effect on customers with low purchase frequency in the pre-subscription period but fails to capture why customers with repeat purchases of known products (i.e., repetitive) are likely to increase their purchases as well.

The core prediction of the fourth explanation (sunk cost fallacy) is that members would increase their purchases by switching their purchases away from competitors, leading to increased share of wallet for the focal company. In order to identify switching behavior, it would be desirable to have complete information on customer purchases from both the focal firm and competitors. Unfortunately, we do not directly observe member purchases from other firms. Instead, we test one prediction for the pattern in switching behavior. If members were to switch
their purchases from competitors to the focal firm, they are more likely to increase their purchases for products with low switching costs (i.e., for products where it is relatively easy to switch brands). We next explain in detail how we test this notion.

We take advantage of data from the survey conducted by a third party to construct two measures of switching costs.\(^\text{16}\) The survey obtains information regarding the extent to which consumers switch brands when shopping for a group of beauty products, called as sub-category in practice, that have similar uses, e.g. foundations, moisturizers, lipsticks, etc.\(^\text{17}\) For example, just over 23\% of consumers never switched their brow pencil, compared to less than 15\% who switched brow pencils regularly, and 32.8\% of customers switched brands for lipsticks all the time while only 12.7\% of customers switched brands for serum. The self-stated information on switching behavior allows us to construct two measures of switching cost, namely, score and rank. First, we calculate the percentage of survey respondents who state that they never switch, less the percentage of respondents who switch all the time (score). Second, we use the descending rank of the score measure to test the robustness of our results (rank).

To enhance the comparison between the data from a third-party survey and our particular context, we use product data that includes information on the product classification both at the product category level (e.g., skin care) and at the sub-category level (e.g., foundations, moisturizers etc. within skin care), and examine the hypothesis that members increase their purchases more so for products in sub-categories with lower switching costs. In the analysis,

\(^{16}\) The survey was conducted by Corra.com, a global digital agency focusing on beauty, fashion, and lifestyle brands. The data from the survey is available at: https://corra.com/beauty-product-loyalty/ (last accessed December 18, 2018).

\(^{17}\) Beauty products in a general product category (e.g., skin care, make-up) are further classified into sub-categories within each category. For example, skin care products are classified further by how they work and what they do, e.g., moisturizers, cleansers, sun care, etc. Make-up products are classified further into foundations, lip sticks, eye shadow, etc. This classification of products to category and sub-category is standard for beauty products and is widely used to categorize beauty products for online commerce. The survey contains information on brand-switching behavior among 13 sub-categories across product categories.
therefore, we restrict our attention to those sub-categories where we can specify a set of measures for switching costs based on the third-party survey.

We first estimate the changes in customer purchases (both in level and in percentage) in sub-categories for each member. Then we compare the treatment effects between sub-categories with low and high switching costs by estimating the following fixed effects model:

\[ \hat{\tau}_{ic} = \beta_0 + \beta_1 \text{Switch}_c + \alpha_i + \epsilon_{ic}, \]

where \( \hat{\tau}_{ic} \) is the estimated treatment effects for customer \( i \) in sub-category \( c \). The variable \( \text{Switch}_c \) is a measure of switching costs (score or rank) in sub-category \( c \). We also include individual-level fixed effects. The coefficient of interest is \( \beta_1 \), which reflects how switching costs moderate customer purchases.

Table 5 reports the results. As shown in the first two columns (in level) of the table, we find that \( \beta_1 \) is negative using either measure of switching cost, suggesting that customers are more likely to increase their purchases in sub-categories where switching costs are lower. The last two columns (in percentage) of the table indicate that the treatment effect is larger for sub-categories with lower switching costs, implying that it is less of a concern that the increase in customer purchases is larger because these products tend to have higher prices. These findings collectively support the notion that the increase in customer purchases is related to customer switching from competitors and contributes to the increased share of wallet for the focal firm.

Insert Table 5 about here

6. Robustness Checks

6.1 Other Cohorts
We replicate our analysis with the members who joined the program in other months. The results suggest that their purchase patterns are qualitatively similar to members used in the main analysis. For instance, members who joined in May and June 2016 increase customer purchases by $25.20 (95% interval [8.21, 57.98]) and $27.79 (95% interval [9.52, 53.20]) per month, respectively.

6.2 Unconfoundedness Assumption

The causal forest framework has advantages over other causal inference methods, e.g., propensity score matching and nearest neighbor matching, in that it matches members and non-members in a non-parametric and robust manner. Since the treatment is not assigned randomly, however, the validity of the method still hinges on the assumption of (conditional) unconfoundedness. That is, conditional on the observables, the treatment status is independent of the potential outcomes. While the unconfoundedness assumption is usually not directly testable, we present two pieces of evidence to alleviate the concern for this assumption.

First, we conduct a balance test to examine whether the forest is successful in achieving comparability between the treatment and the matched control units. The idea behind the test is that since the cohort of members for the main analysis adopted the program in April 2016, their purchase behavior should be similar to that of matched non-members before April 2016 if there is no longer selection. As we have already used the 12-month pre-subscription data (April 2015 to March 2016) to build the forest, we compare the customer purchases prior to that period between members and (matched) non-members. The results suggest that purchase amount in February 2015 is indistinguishable between the two groups (difference = -7.56, 95% interval [-17.13, 3.58]).\footnote{We also conduct the analysis with January 2015 and March 2015, and find qualitatively similar results.}
Second, we use late adopters, rather than non-members, as controls for early adopters. The late adopters could be better controls for the early adopters than non-members if their adoption time is close enough (Manchanda, Packard, and Pattabhiramaiah 2015, Datta, Knox, and Bronnenberg 2017). We choose customers who joined the program between August 2016 and November 2016 as the control group, allowing us to have enough customers to match from and enough time periods to estimate the treatment effect. We find qualitatively similar results: on average, customers increase monthly purchase amount by $31.38 (95% interval [18.23, 55.17]).

7. Conclusions
In this paper, we study the causal impact of customers joining a subscription program on their subsequent purchase behavior. We leverage individual-level transaction data at a company that launched a subscription program at its website. To account for the lack of random assignment in membership status, we use a random classification forest to the change in the outcomes before and after joining the program, match members non-parametrically to non-members with similar characteristics, and compare the changes in purchases within matched pairs. We find that, on average, members increase their purchases by $30 per month, a substantial increase of 150%. The effect of the program on customer purchases is economically significant and persistent. The program keeps customers more engaged as they purchase more often and a larger variety of products. Our main results are robust to different subsamples of the treatment group and other definitions of matching samples.

The causal forest method also allows us to do robust inferences for the individual-level treatment effects. Leveraging heterogeneous treatment effects estimated with the data on program usage, we find that the economic benefits of the subscription program, i.e., unlimited
free shipping and member-exclusive offers (e.g., gift cards, member-exclusive discounts, etc.),
combined do not fully explain the large increase in customer purchases. Nor does channel
substitution explain the changes of purchase behavior at the online channel. While there can be
multiple drivers at work, we suggest that members experience sunk cost fallacy and switch their
purchases away from competitors to the focal company. We provide empirical evidence
consistent with this proposed underlying mechanism as the boost in sales is largest in product
categories with lower switching costs.

Our results shed light on the practice of subscription-based businesses for customer
retention and development. As companies are increasingly concerned about customer
engagement (e.g., breadth and depth of purchases in products and categories), it is important to
understand how subscription programs affect future sales. In our context, we find that such
programs are broadly effective in boosting sales especially by making customers more frequent
shoppers not only in the known categories but also in new categories. While we study a
subscription program launched by a single firm, the insight that the effects of the program on
customer purchases go beyond the tangible economic benefits it contains is not limited to the
structure of our focal program per se. Our results also enable us to derive recommendations
related to the operation of subscription programs. For example, given that customers may be
responding to initial payment, companies can make the initial payment even more salient after
customers become members of the program. Our suggestion is in line with past work that making
prices salient can make members consume a service on a more consistent basis (e.g., Gourville
and Soman 2002).

Since this research is among the first attempts to identify the causal effect of subscription
in retail on customer behavior, naturally, there are several limitations that should be
acknowledged and perhaps addressed in future research. First, we have studied the causal effect of a subscription program for beauty products and it is likely that some of our findings could be reflective of the product category. In line with this notion, the subscription program examined in this research is an online-only program. This choice by the firm in which members benefits from the online channel only, in contrast to other types of subscriptions in which members benefits from both online and offline channels (e.g., Barnes & Noble), has been under much discussion in practice when Amazon bought Whole Foods in 2017. With that in mind, we hope our approach provides a framework for further studies in other product categories (e.g., consumer software, food preparation) and across channels not only in retail but also in other industries (e.g., financial services, business solutions).

Second, our rich data on transaction and program usage comes from a single company. While this is a typical limitation for most empirical work, we are not able to directly observe switching behavior across brands among customers. Our approach is to infer brand switching from the changes in the composition of the product categories customers buy. Richer data, including substitution in product categories across firms, could further enhance our understanding. Some recent work in the area of loyalty programs where transaction data is available across firms is an important step in this direction (e.g., Stourn, Bradlow, and Fader 2017).

Third, we study only the case where a single company launched a subscription program. In the long run, with the relevance and popularity of subscription-based businesses, it is quite possible that several companies in the same industry will have their own subscription programs. With competition in play, the effect of a subscription program on customer engagement and purchase remains unclear. One may expect that the effect of a new program may not be as large but there may also be some interactions between firms’ characteristics (e.g., market share) and
subsequent changes in customer behavior. We hope that our work will inspire future studies to deepen our understanding in this nascent and important area of research.
References


<table>
<thead>
<tr>
<th>Variable</th>
<th>Members</th>
<th>Non-Members</th>
<th>Difference</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purchase amount ($)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>04/2015</td>
<td>9.86</td>
<td>6.36</td>
<td>3.50</td>
<td>0.000</td>
</tr>
<tr>
<td>05/2015</td>
<td>9.77</td>
<td>6.68</td>
<td>3.09</td>
<td>0.000</td>
</tr>
<tr>
<td>06/2015</td>
<td>9.81</td>
<td>7.03</td>
<td>2.78</td>
<td>0.000</td>
</tr>
<tr>
<td>07/2015</td>
<td>9.54</td>
<td>5.84</td>
<td>3.70</td>
<td>0.000</td>
</tr>
<tr>
<td>08/2015</td>
<td>10.05</td>
<td>6.33</td>
<td>3.72</td>
<td>0.000</td>
</tr>
<tr>
<td>09/2015</td>
<td>11.17</td>
<td>7.25</td>
<td>3.92</td>
<td>0.000</td>
</tr>
<tr>
<td>10/2015</td>
<td>14.02</td>
<td>9.63</td>
<td>4.40</td>
<td>0.000</td>
</tr>
<tr>
<td>11/2015</td>
<td>11.18</td>
<td>7.27</td>
<td>3.90</td>
<td>0.000</td>
</tr>
<tr>
<td>12/2015</td>
<td>11.33</td>
<td>4.90</td>
<td>6.43</td>
<td>0.000</td>
</tr>
<tr>
<td>01/2016</td>
<td>16.08</td>
<td>5.50</td>
<td>10.57</td>
<td>0.000</td>
</tr>
<tr>
<td>02/2016</td>
<td>14.21</td>
<td>4.93</td>
<td>9.27</td>
<td>0.000</td>
</tr>
<tr>
<td>03/2016</td>
<td>16.81</td>
<td>4.77</td>
<td>12.05</td>
<td>0.000</td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>35.94</td>
<td>32.84</td>
<td>3.09</td>
<td>0.000</td>
</tr>
<tr>
<td>Gender (Female = 1)</td>
<td>0.93</td>
<td>0.94</td>
<td>-0.01</td>
<td>0.162</td>
</tr>
<tr>
<td>Observations</td>
<td>721</td>
<td>13,768</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Table 2: Covariates for the Causal Forest

<table>
<thead>
<tr>
<th>Covariates</th>
<th>Operationalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer-firm relationship</td>
<td></td>
</tr>
<tr>
<td>Tenure</td>
<td>Elapsed time since having online account</td>
</tr>
<tr>
<td>Breadth</td>
<td>Number of unique categories purchased</td>
</tr>
<tr>
<td>Depth</td>
<td>Number of transactions made</td>
</tr>
<tr>
<td>Basket size</td>
<td>Average basket size</td>
</tr>
<tr>
<td>Monthly purchase amount: Category-level</td>
<td>Monthly spend at each product category</td>
</tr>
<tr>
<td>Monthly purchase amount: Variance</td>
<td>Variance of monthly purchase amount</td>
</tr>
<tr>
<td>Psychographics</td>
<td></td>
</tr>
<tr>
<td>Exploratory</td>
<td>Average time taken for the purchase among three new products purchased since the launch</td>
</tr>
<tr>
<td>Repetitive</td>
<td>1 if a customer made repeat purchases of a product more than ten times, 0 otherwise</td>
</tr>
<tr>
<td>Promotional</td>
<td>Average discount rate received for purchases</td>
</tr>
<tr>
<td>Socio-demographics</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Home address</td>
<td>Coordinates of home address</td>
</tr>
</tbody>
</table>
Table 3: Treatment Effects

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>95% Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purchase amount ($)</td>
<td>29.82</td>
<td>[9.59, 82.79]</td>
</tr>
<tr>
<td>Short run (1 month)</td>
<td>46.20</td>
<td>[16.82, 104.52]</td>
</tr>
<tr>
<td>Medium run (2-3 months)</td>
<td>27.34</td>
<td>[12.77, 40.94]</td>
</tr>
<tr>
<td>Longer run (&gt; 3 months)</td>
<td>28.33</td>
<td>[7.76, 89.25]</td>
</tr>
<tr>
<td>Purchase frequency</td>
<td>1.01</td>
<td>[0.54, 1.64]</td>
</tr>
<tr>
<td>Basket size</td>
<td>-4.96</td>
<td>[-10.57, -1.52]</td>
</tr>
<tr>
<td>Variety ($)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purchase amount in known categories</td>
<td>24.25</td>
<td>[9.56, 55.00]</td>
</tr>
<tr>
<td>Purchase amount in new categories</td>
<td>5.57</td>
<td>[3.84, 37.31]</td>
</tr>
<tr>
<td>Purchase amount of known products</td>
<td>9.23</td>
<td>[3.46, 20.21]</td>
</tr>
<tr>
<td>Purchase amount of new products</td>
<td>20.81</td>
<td>[6.54, 52.36]</td>
</tr>
<tr>
<td></td>
<td>Estimate</td>
<td>Std. Err.</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td><strong>Customer-firm relationship</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tenure</td>
<td>-0.0016***</td>
<td>0.0004</td>
</tr>
<tr>
<td>Breadth</td>
<td>1.05</td>
<td>0.76</td>
</tr>
<tr>
<td>Depth</td>
<td>-2.23***</td>
<td>0.71</td>
</tr>
<tr>
<td>Basket size</td>
<td>-0.05</td>
<td>0.03</td>
</tr>
<tr>
<td>Monthly purchase amount</td>
<td>-69.68</td>
<td>48.82</td>
</tr>
<tr>
<td>Variance of monthly purchase amount</td>
<td>0.0155</td>
<td>0.0139</td>
</tr>
<tr>
<td><strong>Psychographics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exploratory</td>
<td>0.0047***</td>
<td>0.0005</td>
</tr>
<tr>
<td>Repetitive</td>
<td>10.61***</td>
<td>1.88</td>
</tr>
<tr>
<td>Promotional</td>
<td>2.72</td>
<td>2.31</td>
</tr>
<tr>
<td><strong>Socio-demographics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.16***</td>
<td>0.07</td>
</tr>
<tr>
<td>Gender</td>
<td>1.35</td>
<td>2.30</td>
</tr>
<tr>
<td>Intercept</td>
<td>-8.47</td>
<td>9.73</td>
</tr>
<tr>
<td>Observations</td>
<td>721</td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.21</td>
<td></td>
</tr>
</tbody>
</table>

Notes: *** p < 0.01; ** p < 0.05; * p < 0.1.
Table 5: Relationship between Changes in Customer Purchase and Switching Cost

<table>
<thead>
<tr>
<th></th>
<th>Purchase Amount</th>
<th>log(Purchase Amount)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Switching cost</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score</td>
<td>-0.039***</td>
<td>-0.012***</td>
</tr>
<tr>
<td></td>
<td>(0.0014)</td>
<td>(0.0002)</td>
</tr>
<tr>
<td>Rank</td>
<td>-0.089***</td>
<td>-0.026***</td>
</tr>
<tr>
<td></td>
<td>(0.0030)</td>
<td>(0.0004)</td>
</tr>
<tr>
<td><strong>Intercept</strong></td>
<td>0.788***</td>
<td>0.244***</td>
</tr>
<tr>
<td></td>
<td>(0.0066)</td>
<td>(0.0009)</td>
</tr>
<tr>
<td><strong>Individual FE</strong></td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>9,373</td>
<td>9,373</td>
</tr>
<tr>
<td><strong>R-squared</strong></td>
<td>0.065</td>
<td>0.088</td>
</tr>
</tbody>
</table>

Notes: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Standard errors appear in parentheses.
Figure 1: Number of Members by Month
Figure 2: Customer Purchases of Members vs. Non-Members
Figure 3: Distribution of the Treatment Effect

![Graph showing the distribution of the treatment effect](image-url)