# Leveraging Loyalty Programs Using Competitor Based Targeting

Hollenbeck, Brett and Taylor, Wayne

UCLA Anderson, SMU Cox

March 2020

# Leveraging Loyalty Programs Using Competitor Based Targeting

Wayne Taylor
Cox School of Business
Southern Methodist University *

Brett Hollenbeck
Anderson School of Management
UCLA †

November 30, 2020

**Abstract**

Loyalty programs (LPs) are widely used by firms but not well understood. These programs provide discounts and perks to loyal customers and are costly to administer, but produce uncertain changes in spending patterns. We use a large and detailed dataset on customer shopping behavior at one of the largest U.S. retailers before and after joining a loyalty program to evaluate how behavior changes. We combine this with detailed spatial data on customer and store locations, including the locations of competing firms. We find significant changes in behavior associated with joining the LP with a large amount of heterogeneity across customers. We find that location relative to competitors is the factor most strongly associated with increases in spending following joining the LP, suggesting that the LP's quantity discounts work primarily through business stealing and not through other demand expansion. We next estimate a set of predictive models to test what variables best predict how spending will change after joining the LP. We use high-dimensional data on spatial relationships between customers, the focal firm's stores, and competing stores as well as customers' historical spending patterns. These models are used to test whether past sales data reflecting customer's vertical value to the firm or spatial data reflecting customer's horizontal vulnerability are more predictive of post-LP spending increases. We show how LASSO regularization estimated on complex spatial relationships are more predictive than are models using past sales data or simpler spatial models. Finally, we show how firms can use customer and competitor location data to substantially increase LP performance through spatially driven segmentation.

*Keywords: Loyalty programs, predictive analytics, spatial models, retail competition, LASSO estimation*

*wjtaylor@mail.smu.edu
†brett.hollenbeck@anderson.ucla.edu

# 1 Introduction

Loyalty programs (LPs) are now prevalent in many industries. The loyalty programs as marketers know them today originated in the airlines in the 1980s and have since permeated industries such as hotels, casinos, retailers, grocery stores, and restaurants, among others. The pervasiveness of loyalty programs is partly due to their flexibility in how they can be structured (e.g., earning rates and rewards) and managed for strategic decisions (e.g., targeting to increase customer engagement).

In spite of the popularity of loyalty programs with firms, their effectiveness at increasing profits has long been subject to debate. This debate has centered on the costs of giving discounts and perks to the most loyal customers, as well as the costs of administering the program itself, and whether these costs are justified by increases in spending by those customers. Loyalty programs have the potential to increase profits by increasing switching costs for existing customers, stealing business from rivals, or through second degree price discrimination. They may also indirectly increase profits by increasing customers psychological perceptions of the firm, by generating customer data that can be used for targeted promotions or CRM, or by exploiting agency issues such as flights booked by business travelers and paid for by their employers (Dreze and Nunes (2008), Roehm et al. (2002), Verhoef (2003), and Shugan (2005)).[1]

Empirical studies of whether loyalty programs actually do increase profits have found mixed results. Verhoef (2003) finds that the effects are positive but very small, DeWulf et al. (2001) finds no support for positive effects of direct mail, Shugan (2005) finds that firms gain short term revenue at the expense of longer term reward payments, and Hartmann and Viard (2008) found no evidence that loyalty programs create switching costs.

In this paper, we use a large and detailed dataset from one of the largest U.S. retailers on customer shopping behavior before and after joining a loyalty program. This loyalty program takes a common form, tiered discounts after a certain level of spending, essentially operating as 2nd degree price discrimination or one particular form of behavior-based pricing (BBP). We therefore analyze the program through the lens of theoretical work in marketing on price discrimination and BBP, most notably Shin and Sudhir (2010). We leverage our data and this framework to answer these research questions: when loyalty programs take this common form, how do consumer spending patterns change after they join? What factors predict who will increase spending after joining? How

[1]For a more complete review, please see Bijmolt et al. (2011), Liu and Yang (2009), and McCall and Voorhees (2010)

can firms use insight from readily available spatial data on competitive structure to improve LP outcomes? Finally, we show how studying this contributes to the understanding of loyalty programs and price discrimination more broadly.

The practice of BBP in the form of rewarding one's best customers with discounted prices is controversial. Shin and Sudhir (2010) showed how two conditions must be met for this type of pricing strategy to be profitable. First, there must exist substantial heterogeneity in customer value. Second, customer preferences must be stochastic. To investigate these conditions empirically in our retail spending data, we combine detailed customer location data with their data on past spending patterns. We also collect data on the retailer's competitor's locations. A crucial aspect of our data is that it results from merging credit card spending data with customer data. We therefore observe spending both before and after a customer joins the LP. We can therefore exploit variation across markets with different levels of competition and types of competitors at a granular level. We then observe how changes in spending associated with joining the LP varies at the customer level for customers facing different local competitive structures. We can observe, for instance, if customers in more isolated markets change their spending or if they merely start receiving discounts on their purchases and how this differs in highly competitive markets.

We use these features to provide two contributions. First, we contribute to the loyalty program literature by adopting the insight from the study of behavior-based pricing (BBP) on the relative importance of customer vertical value and stochastic preferences, represented by horizontal vulnerability to competitors. We show both descriptively and in model form that customer location and specifically the local competitive structure around each customer are key factors for predicting when the LP is profitable. Second, we show how a firm can operationalize this insight by using spatial factors to form customer segments and target the LP based on this segmentation. We demonstrate how in principle this can increase LP profitability substantially. In both these sets of results we use a common insight, that an individual's spatial relationship with competitor stores is an a primary determinant of when spending increases and that the nature of this relationship, while complex, is typically exploitable with freely available data. To our knowledge, this is one of the first papers to explicitly link the competitive structure of a market with the performance of a loyalty program.[2]

We first take advantage of the unique nature of our data and show a large and comprehensive

[2]Another example is Stourm et al. (2017), who study how distances between retail partners within a coalition loyalty program influence reward spillover activity.

set of descriptive results on customer behavior. Our analysis of the location and spending data finds the following results. First, behavior often changes significantly when a customer joins the loyalty program. This change may or may not be caused by the LP itself, the causation could run in the opposite direction for some customers. However, across all customers the aggregate change in profits associated with customers joining the LP is small or even negative. Second, there is wide variation across customers in profitability, as many low-spenders greatly increase their spending when they join but many of retailer's best customers simply gain large discounts on their purchases. Certain types of LP joiners seem especially important, in particular we identify segments of customers who seem to consolidate their purchases at the focal retailer by increasing trip frequency and customers who maintain the same amount of products purchased but upgrade to higher priced products.[3] Notably, the strongest predictor of whether a customer's joining the LP will be associated with higher or lower profits is whether they are located near a competitor store or live in an isolated market with only the focal store present. In fact, all the profit gains associated with the LP are from customers in competitive markets. While this result is intuitive it has not been shown previously in a LP setting and spatial relationships are not currently used by the firm in targeting their marketing. This motivates the use of spatial data that accounts for the local competitive structure around each customer as the basis of a segmentation strategy to increase the LP's effectiveness.

While simple descriptive results suggest spatial competitive structure is the crucial determinant of LP effectiveness, the full relationship is likely to be highly complex. To capture the complex spatial relationship between the customer, focal firm, and competition, we develop a new method for the treatment of spatial data in predicting customer outcomes. To do so we estimate the relationship between the change in spending (conditional a customer on joining the LP) and a very large number of variables on the complex spatial relationships with competitors at the individual level then select the metrics most predictive of the customer behavior of interest. This estimation is essentially a prediction problem with many potential predictive relationships and is therefore well suited to shrinkage type estimators like LASSO regularization.

Most prior work on LP's suffers from only observing spending data on customers in the program.[4] There is a clear selection effect of high spending customers into the LP that significantly complicates

[3]We emphasize that we do not observe spending at competitor stores, as is typical with firm-level data. We therefore do not literally observe consolidation and business stealing effects but infer them from how changes in behavior within customers vary, especially across customers with different spatial characteristics.

[4]See Meyer-Waarden and Benavent (2009) for an exception.

all efforts to measure the effect of LP's on behavior. Even with pre and post-join data there are still selection effects that would complicate an effort to isolate a purely casual effect. In particular, even when measuring the change in spending before and after within an individual customer, the decision to join the LP could coincide with a planned change in spending. We tackle this selection effect by modeling self-selection explicitly in a first stage model and then using this to correct for selection bias in a second stage model of spending. This takes the form of a flexible control function implementation of a Heckman-style selection correction approach. We find that after correcting for self-selection, our LASSO implementation of spatial variables performs dramatically better at predictive tasks than standard treatments of spatial data.

Next, we compare the performance of models using spatial data and past sales data as inputs. Following the insights of Shin and Sudhir (2010), we broadly define these data on spending patterns as capturing *vertical quality* or the value of a customer in terms of their overall demand, and define the data on local competitive structures as capturing *horizontal quality* or how likely it is to shift a customers spending from a competitor's store to the focal store. This estimation serves two purposes. First, it validates the descriptive result that horizontal quality (i.e., spatial competitive structure) is a stronger predictor of LP effectiveness than vertical quality (i.e., past spending patterns or other RFM variables).

Second, we then use the output to show how firms can take advantage of this insight and leverage customer and competitor location data to increase the performance of their LP through spatially driven segmentation rules. A potential use for this is for firms to develop strategic targeting strategies for LP promotion or to avoid promoting the LP to those whose behavior is unlikely to change and would be using the LP for discounts alone, based on their locations relative to the competition. Because location and travel costs form an important part of preferences over retailers, this can be thought of as a strategy for targeting price discounts on observable preference heterogeneity. Our results show that, consistent with the spirit of recent work in customer valuation (e.g., Ascarza (2018)), firms should leverage sources of observed heterogeneity and focus on the customers who are spatially vulnerable with easy access to competitors. While this result seems intuitive, we show empirically that the firm who provided our data does not currently use spatial information in its targeting and promotion decisions.

In addition to contributing to the understanding of loyalty programs, we contribute to recent work that has explored the benefits of geotargeting, where promotion or other marketing activity is a function of a customer's real-time location using mobile data (Luo et al. (2014), Chen et al. (2016)). Fong et al. (2015) and Dube et al. (2017) have also considered "geoconquesting", where the marketing activity focuses on instances when the customer is located near a competitor as opposed to near the focal firm. This literature is limited, but potential gains from geoconquesting are especially likely when the marketing activity is based in part on business stealing, as it is for loyalty programs.

While the previous literature therefore incorporates select spatial components, mostly it does not fully account for the complex customer-store-competitor spatial relationship. In prior work, competitor information is often integrated into models through simple customer-store or store-competitor distance metrics, thereby eliminating the possibility of complex spatial analyses. We find these simple distance metrics perform especially poorly compared to our approach.

This work also adds to the extensive literature related to a customers' share of wallet, as our findings suggest that demand changes at the focal firm come at the expense of decreasing sales at nearby competitors. Prior work has shown that loyalty program membership can increase share of wallet (Leenheer et al. (2007) and Wirtz et al. (2007)). Chen and Steckel (2012) attempt to solve the "incomplete information problem" of only observing behavior at a single firm by focusing on inter-purchase time distributions. Du et al. (2007) recognize that transaction data within the firm is insufficient on its own and augmentation might be necessary to distinguish customers with large total market potential. However, prior research does not fully capitalize on the locations of competitors to better understand potential share of wallet opportunities.

We also contribute to the study of competitive promotions and competitive price discrimination. Price discrimination strategies such as loyalty rewards should never lower profits by a monopolist but in oligopoly settings this is no longer true as firms may face a prisoner's dilemma (Shaffer and Zhang (1995)). Chen et al. (2001) shows that when individual targeting is possible but imperfect, it can soften price competition among competing firms, but as targeting precision increases the prisoner's dilemma reasserts itself. Ultimately then, it is an empirical question whether and when targeted price discrimination can increase profits. Previous work has shown that in practice the benefits of using targeted pricing can be quite high (Rossi et al. (1996), Besanko et al. (2003)). Li et al. (2018) specifically consider competitive price discrimination across markets and show that the profitability of tailoring prices to local markets depends on both the local market structure and competitive intensity across markets. Our analysis also considers local market structure but with price discounts in the form of a loyalty program. We show that in the context of quantity discounts, segmenting customers based on horizontal versus vertical characteristics are differentially profitable depending on the degree of local competition.

The remainder of the paper proceeds as follows. Section 2 describes the data on retail sales histories and competitor and customer locations. Section 3 provides model free evidence that the spatial competitive structure surrounding a customer is the key determinant of LP effectiveness. Section 4 describes a LASSO estimation of the complex interaction between competitive structure and LP effectiveness and then uses this result to evaluate the relative predictive power of horizontal and vertical data and then use this to form a segmentation strategy. Section 5 discusses additional managerial implications and avenues for future research.

# 2 Data

In this section we describe the customer transaction data, competitive location data, and provide an overview of the spatial metrics used to characterize the competitive structure.

## 2.1 POS Transaction Data

The transaction data comes from a Fortune 500 specialty retailer. The retailer specializes in a range of product categories including lumber, electrical, and paint, among others. This point-of-sale (POS) data from the retailer is highly detailed: we observe the full basket of purchases at the SKU level from a random sample of 10,029 customers between March 2012 and March 2014 across a variety of product categories, regardless of transaction method (e.g., cash or credit card). These sum to over 2.4 million SKU level purchases across 897,819 store trips. On average, each customer has about 90 trips across nearly five different store locations through the two year observation period, spending about \$110 per store visit, \$684 per month (at the firm level, potentially across multiple store locations), and travels about 9.2 kilometers (5.7 miles). The large levels of monthly and per-trip spending are consistent with the nature of the categories sold at this retailer, which are relatively high-priced categories. In addition, many of the retailer's customers are professional customers whose purchases are related to their jobs. This includes many of the members of the loyalty program.[5]

| | | |
|---|---|---|
| Customers | 10,029 | |
| Date range | March 2012 to 2014 | |
| | Mean | SD |
| Trips/customer | 90 | 77 |
| Spend/trip (net, after discounts) | \$110 | \$102 |
| Spend/month (net, after discounts) | \$684 | \$714 |
| Store locations visited | 4.7 | 3.2 |
| Distance by household (km) | 9.2 | 6.2 |

Table 1: Summary Statistics

Crucially for this analysis, the firm providing the data uses a variety of methods to associate transactions with an individual customer (for example, matching addresses across credit cards used in multiple transactions). This allows the firm to observe all transaction activity regardless of loyalty program enrollment. Specifically, it allows us to observe changes in customer behavior at the firm upon joining the LP.[6] We also observe all marketing activity for these customers through the firm's email campaigns: over 900,000 emails were sent to about 39% of the customers and 11% of the customers received promotions specifically encouraging enrollment in the loyalty program.

Importantly, the data also contains the customer's zip code and the latitude/longitude of each of the firm's store locations, which provides us with a complete picture of all customer interactions with the firm across different store locations as well as the specific products purchased at each store location over time.

We further take advantage of a unique feature of this loyalty program in that the program discounts are earned and applied on only a single large category of the firm's goods. We label this a *limited* loyalty program to emphasize the distinct structure. As requested by the firm providing the data, we cannot disclose the category of goods for which the discount applies. Instead, we refer

[5]This is typical of loyalty programs in many settings, for instance an outsize share of members of airline and hotel loyalty programs are professionals who travel frequently for work.

[6]For the average joiner, we observe about 7 months of transaction activity prior to enrollment and about 8.5 months of transactions after joining.

to it as the *focal* category of interest.

The structure of the limited loyalty program is that of a second degree price discrimination or quantity discounts. Customers who reach spending thresholds in the focal category receive discounts on future purchases within that category. The loyalty program consists of three such thresholds, with increasing discounts upon reaching each threshold. Customers who reach a given threshold retain their status until the end of the calendar year, at which point they begin the earning process again. The loyalty program is offered at the firm level, not the store level, so the thresholds can be attained from purchases across multiple locations.[7]

This unique feature of offering a discount only within a focal category provides an additional layer with which to study the interaction between competitive structure and LP effectiveness. This is because the firm competes with both generalists, who also sell across a large variety of categories, and specialists, who only sell the products in the specific category the LP applies. It also allows us to study whether and when the loyalty program purchases are associated with spillovers into other category purchases.

## 2.2 Competitor Data

We augment the focal firm's POS transaction data with location information (i.e., latitude and longitude) from four competitors. We recognize there may be other competitors in a given market however the firm providing the data explicitly stated these four competitors as their primary concern and others as inconsequential.[8]

The four competitors (see Table 2) vary in both size and product breadth. Competitor #1 is a big box store with a wide product variety (BB). The remaining three competitors are small-box stores. Competitor #2 and competitor #3 offer a wide assortment of products (SB1 and SB2), whereas competitor #4 specializes in the focal product category for which the limited loyalty program applies (SS). None of these competitors offered a loyalty program similar in structure to that of the focal firm throughout the duration of our data.

[7]This loyalty program structure is consistent with that of a "customer-tier" LP, as opposed to a "frequency-based" LP, following Blattberg (2008).

[8]This industry is highly concentrated, with the focal firm and the primary big-box competitor capturing about 85% of market share, according to a 2019 industry report by IBISWorld.

| Competitor | Reference | Footprint | Product Breadth |
|---|---|---|---|
| #1 | BB | Big-Box | Wide |
| #2 | SB1 | Small-Box | Wide |
| #3 | SB2 | Small-Box | Wide |
| #4 | SS | Small-Box | Specialized |

Table 2: Competitor Types

The variety in the size and product breadth allows us to compare how the impact of the competitive structure might be competitor or type specific. More importantly, it allows us to gain insight into the extent of category-level versus store-level business stealing.

The latitude and longitude locations of each competitor are collected from the Google Maps API. We first pulled all competitors within a 50 kilometer radius of each of the focal stores, and then pulled all competitors within 50 kilometers of each customer located within a 50 kilometer mile radius of each focal store.[9] This expanded footprint ensures that our definition of the competitive structure is customer centric rather than limited to the perspective of the focal store.

## 2.3 Quantifying the Competitive Structure

Our analysis measures and highlights the impact of the complex spatial relationship between the customer, the focal store, and the competitors on customer behavior and ultimately firm profits. Quantifying this relationship in such a way to accurately reflect the tradeoffs that an individual likely encounters when deciding which store to visit requires several complex considerations.

A common approach in quantifying competitive structure is to simply use the distance between a focal store and the customer along with the distance between the competitor and the focal store (or more commonly still, an indicator variable if they are both within, say, a 5 kilometer radius of the focal store). The drawback of this approach is that it does not jointly consider the customer-store-competitor location, resulting in a potential homogenization of very distinct competitive structures.

This limitation is illustrated in Figure 1: three competitors, $C_1$, $C_2$, and $C_3$, and a customer, $I$,

[9]We selected 50 kilometers because this was the maximum radius allowed by the Google Maps API at the time of pulling this information and exceeds the average distance traveled by household by a factor of about 7.

are positioned near the focal store $S$. Both $C_1$ and $C_2$ are nearly the same distance to the store, but their respective relationships to the customer and the focal store are considerably different. The customer has to pass by the focal store in order to visit the first competitor, $C_1$, which suggests some spatial advantage for the focal store. However, the second competitor, $C_2$ is positioned right next to the customer, acting as a convenient alternative to the focal store. In a different case, $C_2$ and $C_3$ are both equidistant to the customer but one is between the customer and the store and the other is in the opposite direction. One goal of this analysis is to incorporate these complex spatial structures into the manager decision process, a strategy that has heretofore been ignored in the analyses of loyalty program effectiveness.

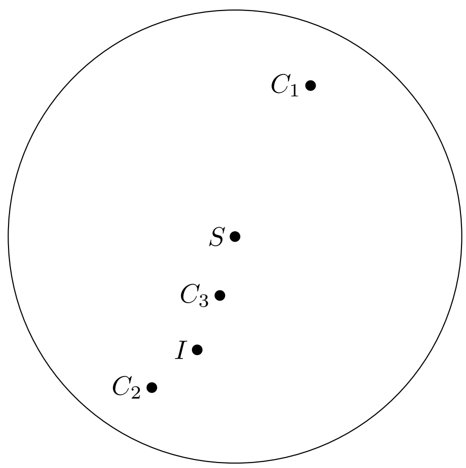



Figure 1: Limitation of Radii Approach

We recognize that it is unreasonable to expect a single metric to capture the complex nature of the spatial relationship between the customer, focal store, and competitor in its entirety. Instead, we approach the problem by proposing a large number of metrics, each of which captures at least some of the complex spatial relationship on its own, and then in our empirical analysis uncover which metrics or which interactions between metrics best predict customer behavior. By honing in on the right combination of distance metrics we can determine which *specific* features of the spatial relationship influence customer behavior the most.

We therefore consider standard distance metrics between each customer and competitor with the the focal firm in addition to the following:

- How much closer is the focal store to the customer, relative to the competitor?
- How *sparse* is the focal store and competitor, relative to the customer?
- Are the focal store and competitor *in the same direction* from the customer, and to what extent?

We also allow for interactions between population density and distance metrics to incorporate the difference in transportation costs between urban and rural areas. For brevity, the complete description of the distance metrics considered are contained in the appendix.

# 3 Spatial Competitive Structure and LP Performance

In this section we take advantage of the unique nature of our data and provide descriptive analysis of how customer behavior interacts with joining a loyalty program. Because we observe sales patterns before and after the customer joins the LP we can measure if this behavior changes and if so, how it changes. These results relate to the extant debates on how and whether LPs are effective at increasing profits.

We then provide model free evidence that the spatial competitive structure are key determinants of LP performance. Our analysis focuses on two metrics of LP effectiveness: the probability that a customer joins the program, and the change in average monthly spend for customers who decide to join. Here and throughout the paper, we use the change in spending net of the LP discounts. For customers who join the LP but do not change shopping patterns, this change is negative by default because of the discounts, and all else equal a positive change for this metric can therefore be taken as strictly beneficial for the retailer. In this section we look at how these metrics relate to the distance between the customer and the focal store and the four competitor types.

One challenge in analyzing LP effectiveness is that the customer's decision to join the loyalty program may be related to unobserved heterogeneity. While we observe purchases both before and after each customer joins the LP, which allows us to condition on individual-level time-invariant factors, it is still true that if customers join due to anticipated changes in their level of spending, the observed change in behavior can not all be attributed to the loyalty program. We therefore do not treat all changes in spending or other behavior associated with joining the LP as having been caused by the LP. Instead we provide descriptive results on how behavior changes and note when those changes in behavior that coincide with joining the LP may be of direct interest even without a purely causal interpretation.

In addition, we treat the location of each customer as essentially exogenous prior to joining the LP, in which case the relative difference in outcomes across customers with different spatial characteristics can provide valid and useful comparisons. Finally, we replicate our results using only spending that does not qualify for the LP discounts and is thus less likely to result from selection on unobservables.

**Descriptive Analysis:**

Table 3 first presents the monthly sales at the firm level, split into qualified and non-qualified spend, for all customers and those that enroll in the loyalty program. In terms of overall spending and spending in the LP category, joiners and non-joiners are actually quite similar. As expected, customers who join the loyalty on average have slightly higher monthly spending on products that qualify for the LP discounts but not higher overall spending, and the differences are modest.

| | All Customers | LP Joiners |
|---|---|---|
| Sales per Month | \$684 | \$662 |
| Qualified Sales per Month | \$31 | \$47 |
| Non-Qualified Sales per Month | \$652 | \$615 |

Table 3: Overall Monthly Spending

We next show the difference between purchase behavior before and after a customer joins the LP. Table 4 presents the average change in sales, trip frequency, and basket level metrics for customers who join the loyalty program.[10] On average, customers tend to increase monthly spending by about \$47 upon joining the loyalty program, with the majority of this change attributed to spend in categories that qualify for LP discounts. This is not surprising, but we also see a positive change in non-qualified spend, which could suggesting that the increase in spending associated with joining the LP spills over into other categories.

Joiners show a small increase in trip frequency, almost no change in number of categories shopped

[10]For all tables presented in the descriptive analysis, we exclude households that fall within the upper or lower .5% of monthly sales prior to joining the loyalty program to prevent the undue influence of outliers for the summary statistics.

in, and actually slightly decrease the number of unique products purchased. We also note there is substantial heterogeneity across customers.

In the second column, we limit attention to customers whose change in monthly spend was positive. Interestingly, for these customers, the vast majority of the increase is attributed to spend outside of the focal category and does not qualify for the LP discounts. Much of this is driven by a large increase in trip frequency, with more modest changes in basket diversity (distinct categories and SKUs), and basket size. We do not suggest that the LP alone caused all of these changes, but note that within the base of customers that do join the loyalty program, there is substantial heterogeneity in post-join behavior.

| | All Joiners | $> 0$ $\Delta$ Sales |
|---|---|---|
| $\Delta$ Monthly Sales | \$47 | \$821 |
| $\Delta$ Monthly Qualified Sales | \$37 | \$133 |
| $\Delta$ Non-Qualified Sales | \$10 | \$688 |
| $\Delta$ Trip Frequency | 0.07 | 3.95 |
| $\Delta$ Unique Categories | -0.01 | 0.14 |
| $\Delta$ Unique SKUs | -0.30 | 0.41 |
| $\Delta$ # Items | -2.44 | 1.22 |

Table 4: Change in Behavior

We next consider how a rich but often overlooked source of variation across customers, the competitive structure, can help in attributing the variation in behavior for customers who join the loyalty program. To start, we first label customers as "isolated" if they are located more than eight kilometers (five miles) away from any of the focal store's competitors and "competitive" otherwise, which indicates there is a competitor relatively close to the customer. We selected this cutoff based on average distance driven by joining customers to surrounding focal stores. Table 5 summarizes a few metrics of interest across these types of customers, namely the probability of joining and the average changes in monthly qualified and non-qualified spend. We also include the average customer sales per month to alleviate concerns that our region labels are associated with substantially different customer types, at least in terms baseline spending patterns.

Approximately one quarter of the customers are located in "competitive" regions and are relatively close to one of the four competitors of interest, as identified by the firm providing the data. Relative to customers in more isolated regions, these customers have essentially the same probability of joining the loyalty program but exhibit a substantially larger change in monthly spend. As expected, the change in qualified spend is positive across both groups. However, customers located near the competition also exhibit an equally high change in non-qualified spend. For these customers, joining the loyalty program may present an opportunity to consolidate purchase activity at the focal store. For customers in relatively isolated areas, they are likely devoting most of their budget to the focal store prior to joining the LP and enrollment is unlikely to generate additional spend from competition.

| | **Competitive** | **Isolated** |
|---|---|---|
| Share of Customers | 25% | 75% |
| Pr(Join) | 4.4% | 4.5% |
| $\Delta$ Sales\|Join | \$109 | \$28 |
| $\Delta$ Qualified Sales\|Join | \$52 | \$33 |
| $\Delta$ Non-Qualified Sales\|Join | \$57 | -\$5 |
| Sales per Month | \$707 | \$647 |

Table 5: Effect of Competition

In Table 6 we further split the change in monthly sales based on the median qualified spending amounts prior to joining the loyalty program. In both cases, the larger change in spend comes from below-median spenders pre-LP. This is somewhat counterintuitive since the program targets discounts at high spenders. But most notably, this change is more than twice as large in competitive regions, suggesting that customers with relatively low spend prior to joining are likely splitting their purchase behavior across competitors. For customers with relatively high pre-join spend, the change is not as large, but still higher in the competitive regions. High spenders in isolated markets actually decrease spending somewhat, once discounts are taken into account. This suggests that the degree of competition near a customer can have a substantial impact on customer engagement with the loyalty program.

| | Competitive | | Isolated | |
|---|---|---|---|---|
| Pre LP Qualified Spend: | Low | High | Low | High |
| $\Delta$ Sales$\mid$Join | \$197 | \$33 | \$82 | -\$23 |

Table 6: Interaction Between Competition and Pre-Join Spending

To further analyze what is driving these findings we break down the results by competitor type. Figure 2 shows the average change in monthly spending based on the relative isolation of the customer.[11] The findings are relatively consistent with the aggregate results, with the exception of a notably larger increase from customers near the big-box generalist, suggesting this type of competitor presents the most opportunities from which to steal business from, and a notably smaller increase from the small-box specialist, who presents the fewest opportunities from which to steal business from.

To better understand this mechanism, recall that the LP only applies to a specific category, not all products, and that one of the competitors (SS) specializes in selling only that focal category. We thus note the portion of change in spend that occurs in the LP category of interest. Figure 2 highlights this by distinguishing the change in spend as either qualifying for the LP discount (within the focal category) or not, where the combined change is noted by a black square. As expected, a large portion of the change in spend is driven by purchases that qualify for LP discounts. However, there is substantial positive changes outside of the focal category. If there was no spillover effects of the loyalty program, we would expect zero change in purchases that will not impact the LP rewards. Upon joining the LP, customers may decide to consolidate purchase behavior with one store rather than cherry picking rewards from the focal store (in the qualifying category) and continuing with the same purchase patterns at other stores in the non-qualifying categories. For customers located in relatively isolated regions, we see that the gains from focal category purchases tend to be offset by declines in other purchase categories, regardless of competitor type.

[11] Note that not all customers may have access to all competitors within a market. Because of these, some customers will not be associated with certain competitors.

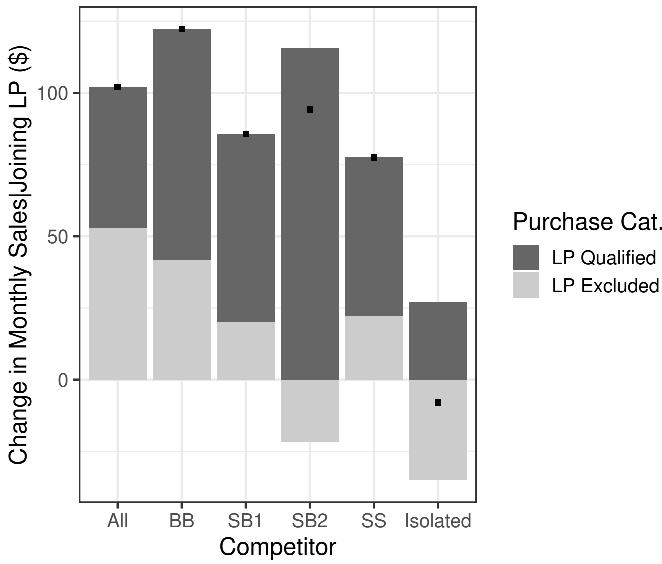



Figure 2: Change in Monthly Spend by Competitor Type and Purchase Category

**Marketing Actions:**

Before continuing with our descriptive analysis, we briefly provide additional information to allay concerns that our findings are simply a reflection of the firm's marketing actions.[12] If the firm were conducting location based targeting during our data time frame, observed differences in spending patterns may be correlated with marketing activity. As previously mentioned, the data contains information on emails sent to each customer. The emails tend to be generic in nature, usually highlighting store-wide sales or new products. We show in the appendix that marketing activity is correlated with spending activity, but found no evidence that email activity is targeted based on location. As our findings show later, this suggests that the focal firm may be mis-targeting its spending by focusing solely on spending differences rather than targeting based on customer location and competitive structure.

In addition to verifying that marketing is not determined by the location of a customer, we also confirmed that the marketing activity does not change for customers upon joining the LP. If this were true, changes in spend may be capturing the change in marketing activity rather than LP enrollment alone. Figure 3 below shows the average email frequency of the customers by week by customer type (joiners and non-joiners). There is clear overlap between groups, highlighting that

[12]Additional details on this supplementary analysis are available in the appendix.

email frequency does not vary once a customer joins the LP. We conclude that any changes in spend after joining the LP are not the result of changes in amount of marketing received.

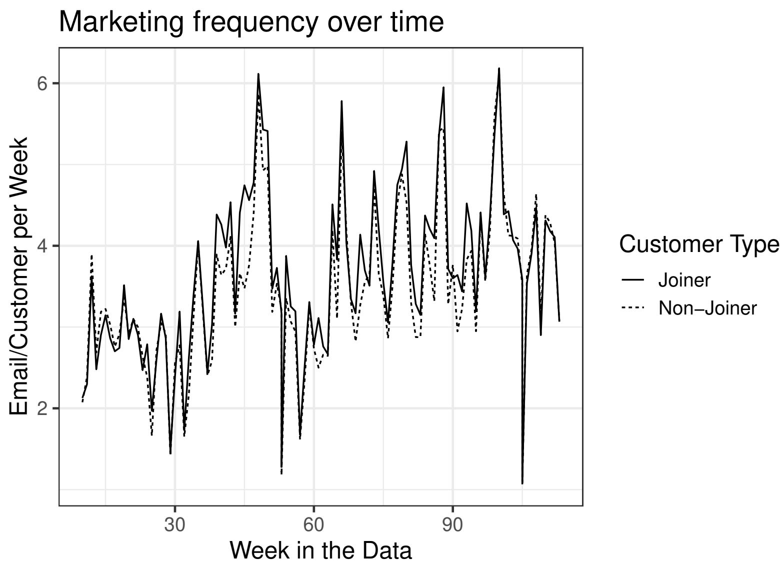



Figure 3: Firm Marketing Activity

**Change in Spend Descriptive Analysis:**

Before specifying our models of join and spend activity, we first provide a deeper analysis of the change in spend based on pre-join customers characteristics. Ideally, the focal firm would be able to use this information to determine which customers are likely to exhibit the largest change in spend after joining the LP and then segment these customers accordingly.

First, Table 7 displays a variety of change metrics based on a median split of five metrics of interest: qualified spending levels prior to joining the LP, number of stores visited, and the three distance metrics of interest: customer-store, customer-competition, and store-competition. We are interested, for instance, in how the firm might predict ex ante what customer characteristics are associated with post-LP increases in spending. Figures 4 and 5 illustrate from the table the change in qualified and non-qualified spend, along with the combined change in spend (shown, as before, as a black square).

Overall, we see the change in sales is higher for customers 1) with relatively low qualified spend prior to joining the LP, 2) who visit more than two stores pre-join, 3) who are relatively far from the focal store, 4) who are relatively close to the competition, and 5) who shop at stores relatively close to a competitor.

Trip frequency increases substantially more for customers with relatively low qualified sales prior to joining the LP, and for customers who are relatively further from the focal store.

The basket composition becomes more diverse (change in unique categories and SKUs) for customers with relatively low qualified sales (pre LP), who visit relatively fewer focal stores, and who are relatively close to the competition. Together, these findings continue to suggest that business stealing is more likely for customers with room to grow with the focal firm but also have access to competitors nearby.

Finally, we redo this analysis but focusing on customers who increase spending after joining the loyalty program. The patterns remain relatively consistent but with the magnitudes more pronounced. As in the section with all joiners, the change in spending (in all categories: overall, qualified, and non-qualified) is substantially higher in situations where the customer is close to focal store and the focal store is close to the competition. A similar pattern holds for changes in trip frequency. Interestingly, when the customer is relatively close to the focal store, the changes are lower than those who are relatively far from the focal store.

These initial results begin to illustrate how the competitive structure, defined as the joint relationship between the customer, focal store, and competitor locations, are strongly related with a customer's engagement with the LP. More importantly, we note that the difference between the high and low categories is substantially smaller when the median split is based on pre-join qualified sales or number of stores visits, suggesting that in some cases spatial characteristics appear to be more informative in identifying change in spend versus traditional behavior based metrics.

**All Joiners**

| | Pre-LP Qual. Sales | | # Stores | | Cust. Store | | Cust. Comp. | | Store Comp. | |
|---|---|---|---|---|---|---|---|---|---|---|
| Median: | $7.57 | | 2 | | 8.44 km | | 14.0 km | | 13.6 km | |
| | low | high | low | high | low | high | low | high | low | high |
| Δ Sales | $109 | -$14 | $28 | $62 | $12 | $82 | $127 | -$32 | $95 | $0 |
| Δ Qual. Sales | $53 | $22 | $41 | $35 | $24 | $51 | $64 | $12 | $46 | $29 |
| Δ Non-Qual. Sales | $56 | -$36 | -$13 | $27 | -$12 | $32 | $63 | -$43 | $49 | -$29 |
| Δ Trip Freq. | 0.37 | -0.23 | 0.11 | 0.04 | -0.13 | 0.26 | 0.13 | 0.01 | 0.02 | 0.11 |
| Δ Unique Cat. | 0.05 | -0.06 | 0.00 | -0.01 | 0.05 | -0.06 | 0.06 | -0.08 | -0.01 | -0.01 |
| Δ Unique SKU | -0.02 | -0.57 | -0.14 | -0.42 | -0.19 | -0.41 | 0.09 | -0.68 | -0.27 | -0.32 |
| Δ # Items | -3.53 | -1.36 | -1.55 | -3.11 | -0.43 | -4.45 | -1.58 | -3.30 | -0.40 | -4.47 |

**Δ Sales|Join > 0**

| | Pre-LP Qual. Sales | | # Stores | | Cust. Store | | Cust. Comp. | | Store Comp. | |
|---|---|---|---|---|---|---|---|---|---|---|
| Median: | $7.75 | | 2 | | 8.24 km | | 14.4 km | | 13.7 km | |
| | low | high | low | high | low | high | low | high | low | high |
| Δ Sales | $911 | $731 | $813 | $826 | $734 | $907 | $960 | $683 | $1,073 | $571 |
| Δ Qual. Sales | $100 | $166 | $116 | $145 | $65 | $200 | $147 | $120 | $157 | $109 |
| Δ Non-Qual. Sales | $812 | $565 | $697 | $681 | $668 | $707 | $813 | $563 | $915 | $462 |
| Δ Trip Freq. | 3.72 | 4.18 | 3.61 | 4.19 | 3.81 | 4.08 | 4.44 | 3.46 | 4.62 | 3.28 |
| Δ Unique Cat. | 0.19 | 0.08 | 0.29 | 0.03 | 0.21 | 0.06 | 0.18 | 0.09 | 0.15 | 0.12 |
| Δ Unique SKU | 0.71 | 0.12 | 0.92 | 0.05 | 0.76 | 0.07 | 0.49 | 0.33 | 0.46 | 0.37 |
| Δ # Items | 1.84 | 0.61 | 4.48 | -1.07 | 5.33 | -2.85 | 4.53 | -2.05 | -1.57 | 3.98 |

Table 7: Change in Behavior Based on Pre-Join Characteristics

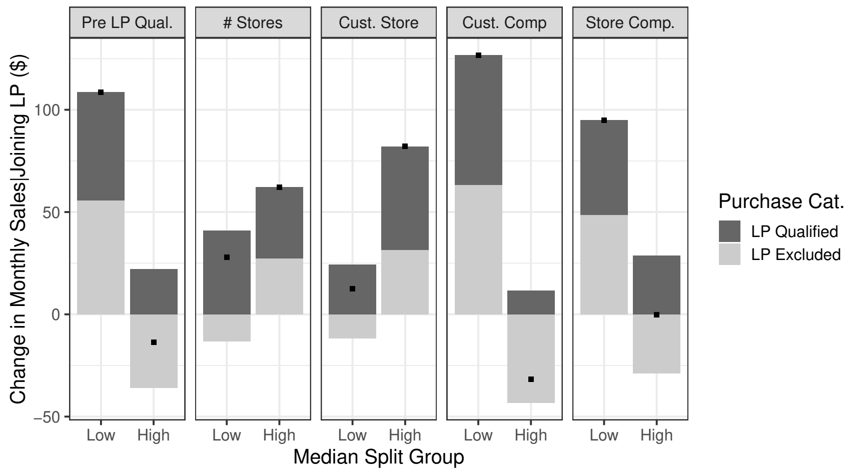



Figure 4: Change in Monthly Spend by Median Split (All Joiners)

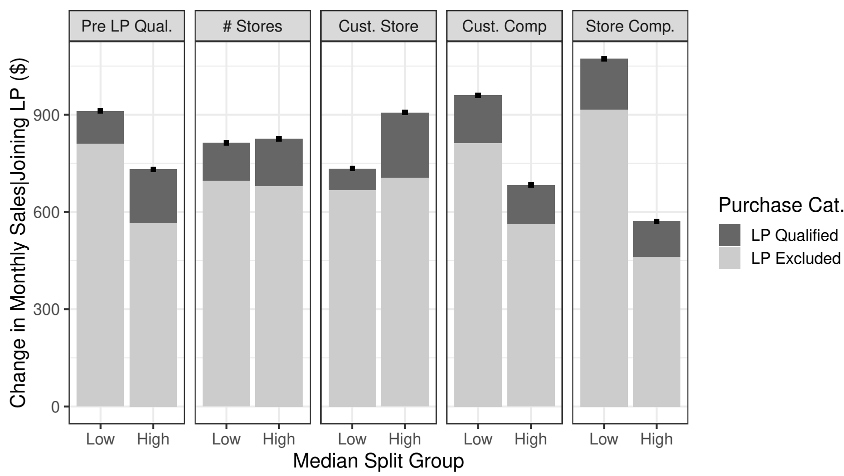



Figure 5: Change in Monthly Spend by Median Split ($\Delta$ Sales|Join $> 0$)

**Patterns Among Profitable Joiners:** Within the customers who increase spending upon joining the LP, we investigate other changes in behavior to determine if there are patterns as to what is driving the positive change in spend. We focus on using the non-spatial variables to see if we might identify segments to provide insight into specifically how engagement with the LP changes upon joining.

In Table 8 we provide a cross-tabulation of selected measures of shopping behavior for these customers. The cells indicate the proportion of customers that fall into each quadrant. We construct the quadrants based on whether there is a substantial increase (which we define as at least a 15% increase in the metric) or if the behavior is flat or even decreasing. For example, of the customers with a positive change in spend, 13.9% of them showed no appreciable increase in trip frequency paired with a significant increase basket size. Within each cell we also note the average increase in spending upon joining to highlight the heterogeneity across customers and relative value of each segment. Figure 6 plots the change in spend figures from this table to more easily compare the cell differences.

As expected, most consumers show an increase in the given variables, since we are selecting only customers who increase spending overall. Of greater interest are the customers who land in the off-diagonal groups - that is, customers who exhibit a positive change in one metric at the expense of a negative change in another. In the complementary Table 9 we provide additional information about the customers within these off-diagonal cells, namely their pre-join behavior and simple spatial metrics.

The first cross tabulation shows that more than one third of those who increase spending do so via increased trip frequency but with no increase in per-trip basket size. This behavior is consistent with consolidating purchases from multiple competing stores to only the focal store after joining the LP and we refer to this group as *consolidators*. The average change in spend of this group at $749 is nearly twice that of the complementary group, who increase basket size but not trip frequency. This is despite the fact that in the complementary table these customers (in the "lower left" cell) have similar pre-join spending levels as those in the opposite cell ("top right"). The spatial information provides some indication of why these customers may have proven to be so valuable. The more valuable group is actually further away from the focal stores, at a distance of more than 10km versus about 8km, and on average but essentially the same distance to the competition. That is, it appears that after joining the customers shift trips that occur elsewhere to the focal store.

The middle cross-tabulation compares changes in number of items purchased and the average price per item. The largest group increases spending by buying more items but without a substantial increase in average price. But about 31% of customers show no increase in basket size but a substantial increase in the price paid per item, suggesting that these customers may be responding to the LP discounts either by upgrading to more expensive items after joining the LP or shifting higher purchases away from competitor stores. We refer to this group as *upgraders*. These customers also increase their total spending by more than twice as much as those who simply buy more items.

We see a similar pattern as the first section: the customers in the more valuable off-diagonal cell (in terms of change in spend after joining) are relatively further away from the focal store. Again, customers that join but are close to the focal store are likely devoting most of their budget to the focal store as is, so joining is unlikely to have a pronounced impact on spend (along with other metrics) unless there is a competitor from which to take business from. The upgraders are also higher spenders pre-join, suggesting they are more likely to receive the LP discounts.

Finally, the bottom cross-tabulation shows again that close to a third of customers have no increase in qualified sales but a substantial increase in non-qualified sales. In terms of change in spend upon joining, those in the top right cell are more than five times more valuable than those in the lower left cell, even though the pre-join sales per month are actually lower. This behavior is again consistent with being a consolidator, shifting non-qualified spending in response to joining the LP. As with before, the spatial metrics provide additional insight into why we might see a marked increase in non-qualified sales: they are more than 2km closer to a competitor, on average, and the focal store is about 1km closer to a competitor, on average.

Our deep analysis of the joining customers continues to provide insight on how behavior changes upon joining the LP and illustrate that spatial metrics provide incremental value in identifying valuable customer groups. In these tables, many of the behavior-based metrics are nearly identical, so any effort to pinpoint the valuable customers based on behavior based metrics alone may be difficult. However, we have provided preliminary evidence that spatial metrics appear to indicate which customers are valuable, both in aggregate analyses as shown earlier and in very detailed analyses shown in this section.

In the next section, we integrate spatial metrics into predictive models of customer behavior. In addition to the relatively simple spatial metrics already presented in the descriptive analysis, we consider a variety of other subtle spatial metrics which may also influence probabilities of joining the loyalty program and change in spend behavior.

| | | $\Delta$ Basket Size (# Items) | |
|---|---|---|---|
| | | Flat/Decrease | Sig. Increase |
| $\Delta$ Trip Frequency | Flat/Decrease | 14.4% | 13.9% |
| | | \$814 | \$312 |
| | Sig. Increase | 35.8% | 35.8% |
| | | \$749 | \$1,094 |
| | | $\Delta$ Price/Item | |
| | | Flat/Decrease | Sig. Increase |
| $\Delta$ Basket Size (# Items) | Flat/Decrease | 19.4% | 30.8% |
| | | \$233 | \$1,103 |
| | Sig. Increase | 34.3% | 15.4% |
| | | \$591 | \$1,506 |
| | | $\Delta$ Non-Qualified Sales | |
| | | Flat/Decrease | Sig. Increase |
| $\Delta$ Qualified Sales | Flat/Decrease | 4.0% | 27.9% |
| | | \$59 | \$636 |
| | Sig. Increase | 12.4% | 55.7% |
| | | \$115 | \$1,125 |

Table 8: Positive $\Delta$ Sales$|$Join Proportions and $\Delta$ Sales

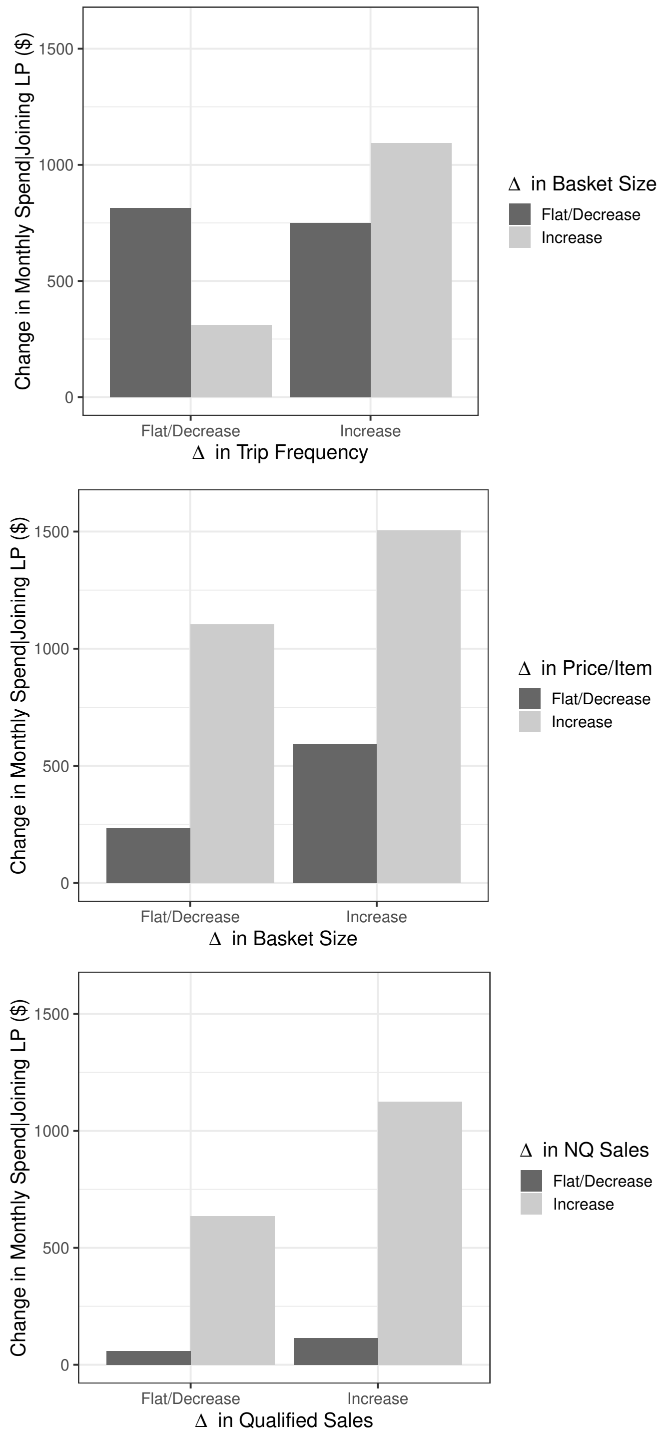



Figure 6: Change in Spend by Metric Cross-Comparison

| | 1: Trip Freq v. # Items | | 2: # Items v. Price/Item | | 3: Qual v. Non-Qual Sales | |
|---|---|---|---|---|---|---|
| Pre LP Metrics | Top Right | Lower Left | Top Right | Lower Left | Top Right | Lower Left |
| Sales per Month | \$506 | \$464 | \$604 | \$414 | \$448 | \$581 |
| Non-Qual Sales/Month | \$471 | \$442 | \$577 | \$388 | \$410 | \$543 |
| Qual Sales/Month | \$35 | \$22 | \$27 | \$26 | \$38 | \$38 |
| Trips/Month | 6.66 | 4.36 | 5.33 | 4.97 | 5.32 | 6.07 |
| Unique Categories | 1.73 | 2.06 | 2.08 | 1.86 | 2.08 | 1.90 |
| Unique SKUs | 3.12 | 4.10 | 4.17 | 3.49 | 4.02 | 3.80 |
| Basket Size | 6.42 | 18.32 | 18.50 | 7.32 | 9.48 | 11.23 |
| Price/Item | \$17.86 | \$10.72 | \$9.88 | \$19.82 | \$11.34 | \$13.36 |
| Stores Visited | 1.96 | 2.29 | 2.21 | 2.06 | 2.04 | 2.44 |
| Store-Customer km | 8.24 | 10.06 | 9.46 | 8.40 | 9.27 | 9.98 |
| Cust.-Competitor km | 15.71 | 16.13 | 15.79 | 13.75 | 14.34 | 16.44 |
| Store-Competitor km | 14.59 | 14.18 | 14.02 | 13.51 | 14.20 | 15.14 |

Table 9: Quadrants of Interest for Positive $\Delta$ Sales|Join

# 4 LASSO Regularization and Competitive Segmentation

In this section we propose and estimate a model of how customer behavior changes after joining the loyalty program. This model takes as inputs detailed data on pre-join spending patterns as well as complex representations of the spatial relationships between customers and their local competitive structures. We broadly define these data on past spending patterns as capturing *vertical quality* or the value of a customer in terms of their overall demand, and define the data on local competitive structures as capturing *horizontal quality* or how likely it is to shift a customers spending from a competitor's store to the focal store. While spending at competitors is unobserved, spatial data can help fill this role. Estimating a model with these two inputs serves two purposes. First, we can analyze the estimation results directly to compare the relative predictive power of our competitive structure (horizontal) variables and traditional predictors of LP effectiveness like sales history (vertical) variables. This relative predictive power is informative on how LPs manage to increase spending, i.e., via demand expansion or business stealing.

Second, the estimation provides a segmentation strategy for a firm seeking to improve LP effectiveness by providing predictions of which customer segments are the most likely to increase spending based on their past spending histories and unique competitive structures. The wide variation across customers' change in spending suggests the gains from segmentation strategies can be quite large.

We face two challenges in this estimation. The first is that the decision to join the loyalty program must be treated as endogenous to change in spending. That is, the customer may join the loyalty program in part because of an anticipated change in spend. To overcome this source of endogeneity, we jointly model the decision to join and the subsequent change in spending and correct for potential selection bias using a control function approach. The second challenge is that representing the nuances of various competitive structures leads to the creation of a very large number of potentially highly correlated variables. To handle this complexity, we employ regularization methods to identify the spatial covariates that are most relevant. Our research objective, along with these challenges, inform the development of our modeling approach.

We therefore estimate a two-stage model, the first stage of which is a model of which customers join the loyalty program as a function of their observable characteristics, including their spending history inside and outside the focal category and their spatial competitive structure. The goal of the first stage is to model customer self-selection into the program in order to account for this in the second stage. The second stage is a predictive model of the change in spending conditional on joining the LP as a function of a customer's spending history and local competitive structure. Because our research objective relies on predictive accuracy (as opposed to inference) and the data are high-dimensional, the problem is well-suited to statistical methods built around dimension reduction or "regularization." We ultimately use a LASSO approach because this will allow us to test inclusion of a large number of possible spatial measures and let the model select the most predictive variables. The output also provides a clearer interpretation than other methods, allowing us to easily identify and assess which individual factors best predict LP effectiveness. In addition, we emphasize that our goal is to not identify the best estimation method for predicting changes in spend, but instead our focus is on introducing novel variables (such as spatial information) into the predictive process and show how even with a simple framework, such as LASSO, the spatial information is valuable. Certainly, the overall predictive power is likely to improve as the methods become more sophisticated, such as with random forests or extreme gradient boosting, but this is outside the scope of our analysis. Finally, the results can be used to assess whether the vertical measures representing spending-based customer value or the horizontal variables representing location-based customer vulnerability to competitors are more effective at predicting behavior. In our discussion of the results we show how this comparison is informative of how and when quantity discounts or behavior-based pricing schemes are likely to succeed.

## 4.1 Two-Stage Model of LP Effects

In this subsection we provide an overview of our estimation method and in the next subsection we discuss the details of how we implement it, including what variables and functional forms are employed.

The decision to join a loyalty program is complex. Customers may join loyalty programs for a variety of reasons, including utilitarian (e.g., discounts), hedonic (e.g., personalized treatment), and symbolic (e.g., social status) (Bijmolt et al. (2011)). We expect this decision to be based on both observed characteristics, such as distance to the store, as well as unobserved characteristics, such as anticipated changes in behavior. This self-selection into the program could generate biased predictions of how spending changes because joiners have different unobserved shocks than non-joiners.

Our empirical strategy is built on a predictive two-stage approach that is modified to account for the high-dimensional spatial information. The model consists of a first stage join decision, where we model the indirect utility of joining the LP as:

$$u_{is} = f(X_{1is}, X_{2is}, Z_{is}) + \mu_{is} + \eta_{is} \tag{1}$$

We include three types of variables in the join decision, $X_1$ contains the spatial competitive structure parameters between customer $i$ and focal store location $s$ and $X_2$ contains firm-level marketing and store-specific customer purchase behavior. We also include several additional variables in $Z$ that can be used to increase the identifying power of the selection correction in the second stage model. We describe each of these in greater detail below. The error term $\eta_{is}$ is an idiosyncratic shock and $\mu_{is}$ is an additional shock that we interpret as a private information shock related to planned spending known to the customer at the time they decide whether or not to join the LP.

Both are IID and mean zero. The first stage results in a binary choice of join or do not join as a function of this indirect utility.

In the second stage, we model the change in spending as:

$$y_{is} = g_1(X_{1is}, X_{2is}) + g_2(\mu_{is}) + \varepsilon_{is} \tag{2}$$

Here $\varepsilon_{is}$ is an unobserved, normally distributed random error centered at zero and represents idiosyncratic shocks to spending and $\mu_{is}$ is the same shock from equation 1 that represents the source of selection bias. The functions $g_1(\cdot)$ and $g_2(\cdot)$ will be described in detail below. We again split the observable characteristics into two sets to highlight the fact that our goal is evaluate the relative predictive power of two types of inputs to our model. $X_1$ contains the spatial relationship between the customer and nearby stores and $X_2$ contains firm-level marketing and store-specific customer purchase behavior.

The purpose of the first stage model is to simply produce a consistent estimate of the probability of joining which can be entered into the second stage as a control function. If the predicted probability of joining is $\hat{P}_{is}(X_{1is}, X_{2is}, Z_{is})$, we can re-write equation 2 as:

$$y_{is} = g_1(X_{1is}, X_{2is}) + \Lambda(\hat{P}_{is}) + \varepsilon_{is} \tag{3}$$

In the classic Heckman (1979) classic selection model, the first stage is estimated as a probit in which case $\Lambda(\hat{P}_{is})$ takes the form of the inverse Mills ratio $\lambda(\hat{f}(X_{1is}, X_{2is}, Z_{is}))$. Without this correction, if $\mu_{is}$ is correlated with $\varepsilon_{is}$ the coefficients of $g_1(\cdot)$ could be biased. While this standard formulation can be estimated using the same set of observed characteristics $(X_1, X_2)$ in the first and second stages, i.e. with no excluded variables, this approach has been found to suffer from high collinearity and to produce very large standard errors (see Wooldridge (2002) chapter 17.4.1). We therefore include a set of first stage variables $Z_{is}$ that influence join probabilities but do not influence the second stage outcome $y_{is}$ and are excluded from the second stage model.[13] These are described in more detail below.

We also generalize the basic model by taking a control function approach: see Heckman and Robb Jr (1985); Heckman and Robb (1986) and Ahn and Powell (1993) for an introduction and

[13]An alternative method that also does not require excluded variables would be to estimate the two-stage model in a Bayesian framework, as in Ghose and Yang (2009). For a full discussion of the identification and use of models of marketing actions with sample selection, see Wachtel and Otter (2013).

extensions to semi-parametric control functions. As in Ellickson and Misra (2012), we let $\Lambda(\cdot)$ take polynomials of $\hat{P}_{is}$, the main advantage being it does not require parametric assumptions on the error structure such as joint normality.[14] The construction of $\Lambda(\cdot)$ on $\hat{P}_{is}$ should be as flexible as possible since the true functional form can rarely be specified a priori. Intuitively, $\Lambda(\cdot)$ serves a similar purpose as an inverse Mills ratio by correcting potential selection bias in the second stage model.

Next, our method allows that the set of spatial variables $X_1$ is a potentially high dimensional object and may also contain variables that are highly correlated with each other. Because of the problems this can cause in selection models (Leung and Yu (1996)) we use regularization when estimating both $g_1(X_{1is}, X_{2is})$ in the second stage and $f(X_{1is}, X_{2is}, Z_{is})$ in the first stage. In both cases we use LASSO regularization. This approach is similar to the widely used Belloni et al. (2014) double-lasso procedure for conducting causal inference with large numbers of variables. Our approach is different in that the first stage LASSO is used in the context of a control function and instead of attempting to conduct inference over a single treatment effect our goal is to make valid predictions of our outcome variable using a large number of explanatory variables.[15] Since both our first stage and second stage goals relate to prediction they are especially well suited to the use of LASSO or similar methods.

First, we run a LASSO regression on the full set of variables $(X_{1is}, X_{2is}, Z_{is})$ to predict the join probability. Then we predict change in spending $y_{is}$ with LASSO regularization on $(X_{1is}, X_{2is})$ after offsetting the control function $\Lambda(\hat{P}_{is})$.[16] Intuitively, this method works by modeling the join decision as a function of observed characteristics in order to recover the extent to which customers have a large or small private information shock to the value of joining. If a consumer has a very high probability of joining according to their observed characteristics, for instance a high level of spending in the focal category, the expected value of the private information shock is low. By contrast, if a consumer has a very low probability of joining according to their observed characteristics the expected value of their private information shock must be high if we observe them joining. This selection correction term therefore accounts for the unobserved determinants of the join decision

[14] For other recent examples of control functions to correct selection bias in marketing, see Jain et al. (2020), Phillips et al. (2015), or Nambiar et al. (2019), which combines a control function estimation with a variant of ridge regression.

[15] In principle, we could run a post-LASSO regression along the lines suggested by Belloni et al. (2014) and conduct inference on the coefficients if this was our goal.

[16] This method is similar to Xu et al. (2020) and Cook and Siddiqui (2020) (who use Random Forest instead of LASSO).

in the second stage change in spending. This approach will thereby eliminate or reduce any bias induced by self-section and improve the model's predictive accuracy.

## 4.2 Model Implementation

We split the second stage observable characteristics into two sets to highlight the fact that our goal is evaluate the relative predictive power of two types of inputs to our model. $X_{1is}$ contains the spatial competitive structure metrics between customer $i$ at focal store location $s$ and $X_{2is}$ contains firm-level marketing and store-specific customer purchase behavior. These metrics are intended to jointly capture the complex spatial relationships between the customer, focal store, and competitors. Since there can be more than one focal store in the vicinity of an individual, these metrics vary across each store $s$ for a given individual $i$. Likewise, since each customer's location is unique, the spatial metrics also vary across each individual $i$ from all observations for store $s$.

For $X_{1is}$ we include traditional spatial metrics such as distance between the focal store and the customer and distance between the customer and each competitor, along with more novel measures such as the angle formed between the focal store and the nearest competitor (intended to reflect whether a change in travel direction is required to switch store purchasing). The full list of the metrics created is shown in the appendix. Our aim is to represent the competitive structure in the most flexible way feasible, and then call upon regularization to identify which features of the competitive structure are most related to changes in spend behavior.

$X_{2is}$ includes traditional recency, frequency, and monetary value (RFM) variables along with basket specific metrics: monthly trip frequency, overall sales, distinct number of items, number of product categories, basket size, discounts received, sales in the category of interest, and sales in the category of interest for those products that are included in the limited LP.

We also include a set of variables $Z_{is}$ that serve to facilitate the selection correction approach by entering the first stage model but are excluded from the second stage model. To be effective, these variables need to be correlated with the decision to join the LP in the first stage, $u_{is}$, but have no direct effect on the change in monthly spend in the second stage. We identified five variables that we argue satisfy this criteria and discuss each in turn.

We posit that a customers' distribution in spend across product categories indicates the likelihood of joining the LP, but does not influence the amount of change in monthly spend. Thus we include the proportion of sales in the LP category of interest. Conditional on all other variables in the model, we expect that customers with a relatively high share of spend in the focal category are more likely to join the LP. However, it is not clear how, say, a high proportion would lead to predictable changes in spend upon joining, for two reasons. First, this metric is indifferent to magnitude of spending, so customers with equally high proportions may be spending at very different levels, which will probably influence how much change in spending is possible. Second, a high proportion does not indicate whether a customer has room to grow in the category (either from business stealing, or anticipated increases in spend) or if the decision to join is a reflection of past, unsustainable levels of purchase behavior in the category, which might result in a negative change in spend.

Next, we argue that marketing actions influence short term individual decisions to join the LP, but do not directly impact longer term changes in average monthly spend. Given this, we include two indicators for marketing activity. The first is an indicator for whether the customer received any marketing email, the second for emails specifically encouraging them to join the LP. The impact of LP related marketing on customers joining the LP should be clear. Non-LP marketing emails may also increase the likelihood of joining the LP by increasing brand awareness. In all of these cases the first stage effects are testable and we discuss them when we estimate the first stage model.

The primary concern with these marketing variables is if they are targeted in a way that makes them correlated with the second stage outcome variable change-in-spend. We argue that while these are likely correlated with pre-join spending levels (although notably not with spatial variables)[17], a large set of these spending-related variables are already included as covariates in the second stage model. Beyond each of these, the firm would need to target on unobserved characteristics related to change-in-spend, which seems unlikely given the results already shown. In addition, while marketing emails may cause short term increases in spending, there is no change in the timing or frequency of these emails after a customer joins the LP and so no reason they would cause a change in overall spending from the pre-join period to the post-join period (as shown in Figure 3).

Finally, we go one step further and posit that the *timing* of the emails that encourage the customer to join the LP are exogenous with respect to changes in spend but has a strong influence the likelihood of joining. We therefore construct as variables the number of days between when these emails were sent and 1) the next purchase in the LP focal category and 2) the next purchase that would qualify for an LP discount. We argue that the timing of the email relative to the next

[17] see Table A.2 and Figure A.3 in the supplementary marketing analysis located in the appendix

purchase should be independent of how much they change spending post-join.

Consider two scenarios. First, if the timing of purchases is exogenous with respect to marketing then this variable is inherently exogenous as well. Yet it will plausibly have a first stage effect if the customer is more inclined to join the LP while making an in-category purchase if they have recently received marketing for the LP. Second, if the timing of purchases and LP joining is directly affected by the LP email then the first stage should be even stronger but in neither of these cases is there reason to think the specific timing of the marketing emails is correlated with the unobserved shock to post-join change in spending.

In the following section we test the first stage power of each of these five excluded variables. In our estimation we test including each of these separately as well as including all five to be transparent about to what extent our results are influenced by each. To increase the statistical power of the selection correction approach our preferred specification is to include each of them.

**LASSO Estimation** Our proposed model contains a large number of spatial variables defined to capture the competitive structure. On the one hand, capturing the spatial relationship between the customer, focal store, and competitors is complex and may require many variables. However, all else being equal a concise model is preferable from a managerial standpoint. In addition, many of the spatial metrics we devised may be redundant or highly correlated with each other. To systematically remove variables that are either unnecessary or redundant we employ the Least Absolute Shrinkage and Selection Operator (LASSO) estimator, introduced by Frank and Friedman (1993) and Tibshirani (1996).

In general, for a model with a $k$-dimensional parameter vector $\theta$ the LASSO method performs the following:

$$(\theta^*) = \arg\min \left\{ -\log L(\theta) + \lambda \sum_k |\theta_k| \right\} \tag{4}$$

Where $L$ is the likelihood of data given the model specified, and $\lambda$ is a tuning parameter that represents the penalty incurred if we choose a nonzero value for any parameter $\theta_k$. This approach regulates the trade-off between an accurate model with more predictive power and a concise model that is more readily interpretable by managers.[18] We select the tuning parameter $\lambda$ through ten

[18]An alternative approach would be a ridge regression or similar method. In this case the penalty is applied more smoothly, shrinking coefficients on highly correlated variables towards each other. We prefer the LASSO approach

fold cross-validation, which is perhaps the simplest and most widely used method for this task.[19]

## 4.3 Estimation Results

The results from the LASSO estimation are presented in Table 10. For model validation purposes, we split the original data where 75% of the customers are placed into the training set and 25% into the test set. The results presented below were estimated using the training set alone, while the test set is used to evaluate predictive performance. For brevity, the table only displays variables that are significant in at least one of the two models.

We begin with a review of the first stage results of the join incidence, with an emphasis on the performance of the exclusion variables. Then, we move to the second stage results of changes in monthly spend.

### 4.3.1 First Stage Results: Exclusion Restrictions and Join Probabilities

The primary purpose of the first stage model is to provide estimates to be used for our control function approach to correct for potential selection bias. Because of this, we focus on the performance of the five variables used to satisfy the exclusion restriction.

The first stage model shows that four of the five exclusion restrictions have been retained by the LASSO regularization, with coefficients moving in the expected directions based on our earlier arguments. First, we see a strong positive effect on the proportion of sales in the category of interest: customers who dedicate a greater portion of store spend to the LP category are more likely to join the LP. We also see evidence that both marketing indicators have a positive coefficient, with a much higher coefficient on the emails specifically encouraging an LP join. Last, we see that that the coefficient on the days between marketing and a purchase in the LP focal category is negative.

It's important to recognize that the typical constructs like $p$-values do not exist for LASSO estimates (Lockhart et al. (2014)), so additional care may be required to measure significance, at least in the usual statistical sense. Because of this, as a robustness check we also estimate the first stage model using a post-LASSO approach, in the style of Belloni et al. (2013). Specifically, we use

because the penalty structure removes coefficients entirely, resulting in clearer model interpretation. Specifically, we can see that many non-spatial variables will drop out altogether. This result is a notable output of interest. A ridge regression would lack this clean interpretation but would result in similar quantitative predictions and can be provided upon request.

[19] See James et al. (2013) for more information.

the LASSO results to identify which variables to then use in a standard logistic regression. The results are presented in Table A.4 and show similar results to LASSO: the same exclusion variable coefficients are significant with nearly identical values.[20] The one exclusion variable that is not significant is the time between a marketing email and the next purchase that would qualify for a LP discount. We suspect this is partially due to its correlation with the other exclusion restriction that measures time between marketing and a purchase in the focal category.

As an additional robustness check, we conduct a similar post-LASSO approach with but only including a single exclusion variable at a time. The results are shown in the last five columns of Table A.4 and highlight that the coefficients on the exclusion restrictions remain stable even on their own. The exception is the last exclusion restriction on the timing between the marketing activity and a LP qualifying purchase, which, in the absence of the similar exclusion restriction, shows a significant negative coefficient, again suggesting that these two variables together may be redundant.

Finally, to confirm that these five exclusion variables together significantly improve the model, we conduct a likelihood ratio test between the full model and a model removing these exclusion variables. We found a test statistic of $\chi^2(5) = 741.05$ resulting in a $p$-value under .0001. The Akaike information criterion (AIC) and Bayesian information criterion (BIC) also support the model with the five exclusion restrictions. Overall, the evidence suggests that our exclusion variables are contributing to the first stage model fit.

### 4.3.2 Second Stage: Change in Monthly Spend

In this section we show the second stage results. We in particular are interested in whether the competitive structure metrics serve as useful predictors of the estimated LP effectiveness, as measured by change in monthly spending at the focal store, conditional on joining the firm's loyalty program. Recall the dependent variable in this estimation is the change in average monthly spend after joining the loyalty program, net of any LP discounts. We focus on average monthly spend to account for different lengths of data before and after the customer joins. In addition, for estimation we only included customers who had at least some spending both before and after joining. Since we are interested in how spending *changes* after joining, rather than *whether* spending occurs or

[20]Due to the regularization method in LASSO, there is a tendency to bias estimates towards zero. In the post-LASSO results, the five exclusion variable coefficients exhibit slightly higher magnitudes than the coefficients from the main LASSO table.

not, this seems to be a reasonable filter.

First, there is a positive impact on the squared distance between the customer and the focal store, suggesting that the greater gains come from those located further away from the focal store. This aligns with our previous intuition: customers located near the focal store are more likely to already dedicate the majority of their spend with the focal store, so the potential for increases in spend are limited, relative to those who are further away and thus more likely to be sharing spend with the competition. The distance between the customer and the competition also influences changes in spend at the focal store. Customers who exhibit the largest changes in spend are those who are relatively far from the the big-box generalist (ic1sq) and the second small-box generalist (ic3sq) and close to the other competitors, holding all other variables constant.

Many of the radius band coefficients drop out of this model. However, we see that the intercept angle is retained. The difference in the estimated change in spend for a customer where a big-box (BB) competitor is in the same line of travel as the focal store (intercept angle of zero) versus in the opposite direction (intercept angle of 180 degrees) is about $122, holding all other variables constant. This is a non-trivial amount that cannot be captured using traditional spatial measures. However, it is important to again point out that many of these distance metrics cannot vary independently of others. The visual results presented later will illustrate the degree to which the intercept angle influences the change in spend while simultaneously accounting for other changes.

Finally, population density along with its interactions with the direct distance measures has a significant impact on the change in monthly spend. On its own, the effect is positive: more densely populated areas are associated with a greater change in spend. This is intuitive as more densely populated areas tend to have a greater level of competition and higher incomes. There are also numerous interactions with our simple distance metrics. For instance, increases in the distance between the store and the customer influence customers in dense areas more negatively, relative to customers in less dense areas. This appears to be capturing the challenges of traveling in more densely populated areas (e.g., traveling a given distance in a city versus a rural setting).

As previously argued, the construction of our exclusion variables, along with their estimation results, lend credibility in our attempt to control for selection bias in the second stage variables. We are also able to take advantage of the unique LP design to conduct an additional robustness check. In the results table we include a column where the estimation proceeds as before but only using the change in spending on products that *do not* qualify for the LP discounts. Comparing this column and the column for all spending shows a very high correlation in the LASSO coefficients and in which variables are retained. Since the products retained for the third column do not qualify for the LP discount, there is less reason to expect selection bias to be present. More specifically, there is no incentive for customers to self-select into the LP in the anticipation of spending changes on products that will cost exactly the same even after joining. Given that the coefficients in the last two columns are similar, we are confident that either a) our control function approach is effective or b) self-selection was not a prominent issue in this data. In either case, our second stage coefficients are credible.

### 4.3.3 Summary

The LASSO results suggest that the customer behavior appears to be strongly influenced by the competitive structure of the local market. The regularization tends to keep quite a few of the spatial metrics, suggesting that relatively simple distance metrics are, on their own, unable to fully characterize the predicted behavior.

Table 10: Lasso Pr(Join) and Change in Sales|Join

| | Coefficient | Pr(Join) | Change in Sales\|Join | Change in Non-Qualified Sales\|Join |
|---|---|---|---|---|
| 1 | (Intercept) | -3.5049 | -10.5927 | -44.4793 |
| 2 | sisq | | 0.0038 | 0.0122 |
| 3 | ic2 | 4.5936e-04 | | |
| 4 | ic3 | 0.0054 | | |
| 5 | ic1sq | | 0.2102 | 0.2068 |
| 6 | ic2sq | 3.1927e-05 | -0.1383 | -0.1372 |
| 7 | ic3sq | | 0.1505 | 0.1395 |
| 8 | ic4sq | -8.1218e-05 | -0.1315 | -0.1046 |
| 9 | sc1 | 0.0027 | | |
| 10 | sc1sq | 7.2080e-05 | -0.0173 | -0.0215 |
| 11 | sc2sq | | -0.0552 | -0.0368 |
| 12 | sc3sq | 5.1451e-05 | -0.1101 | -0.1184 |
| 13 | sc4sq | -6.7231e-05 | 0.1154 | 0.0986 |
| 14 | s4 | -0.3108 | | |
| 15 | nAll_1 | 0.0218 | | |
| 16 | n1_1 | 0.6901 | | |
| 17 | n3_1 | 0.1956 | | |
| 18 | n5_1 | 0.1568 | | |
| 19 | n10_1 | -0.0470 | | |
| 20 | n15_1 | 0.0519 | | |
| 21 | nAll_2 | -0.0015 | 0.4675 | 0.4056 |
| 22 | n1_2 | -0.2532 | | |
| 23 | n10_2 | 0.0339 | | |
| 24 | n15_2 | -0.0167 | | |
| 25 | nAll_3 | -0.0141 | | |
| 26 | n1_3 | 0.2896 | | |
| 27 | n3_3 | -0.0029 | | |
| 28 | n5_3 | -0.0703 | | |
| 29 | n10_3 | 0.0582 | | |
| 30 | nAll_4 | -0.0076 | | |
| 31 | n3_4 | 0.1386 | | |
| 32 | n15_4 | -0.0040 | | |
| 33 | cInt1: BB intercept angle | 0.0013 | 0.6764 | 0.7403 |
| 34 | cInt2: SB1 intercept angle | | -0.2323 | -0.2871 |
| 35 | cInt3: SB2 intercept angle | 4.3567e-04 | 0.5183 | 0.5493 |
| 36 | ccbar_1 (avg. BB-BB comp. center km) | -0.0123 | | |
| 37 | ccbar_2 (avg. SB1-SB1 comp. center km) | 0.0091 | 0.9408 | 0.8450 |
| 38 | icbar_2 (avg. customer-SB1 km) | 0.0044 | | |
| 39 | icbar_3 (avg. customer-SB2 km) | 0.0041 | | |
| 40 | icbar_4 (avg. customer-SS km) | -1.1230e-04 | | |
| 41 | Trips/month (pre LP) | 0.0105 | -26.4890 | -25.6314 |
| 42 | Total sales (pre LP) | 1.7945e-04 | 0.0065 | 0.0017 |
| 43 | Total sales in LP category (pre LP) | | 0.0334 | 0.0122 |
| 44 | Total qualified sales in LP category (pre LP) | | -0.1101 | -0.0460 |
| 45 | Number of focal stores w/i 60km | 0.0036 | 1.0282 | 1.0824 |
| 46 | Population Density (in sq miles) | | 0.0279 | 0.0250 |
| 47 | Population Density x si | | -0.0018 | -0.0017 |
| 48 | Population Density x ic1 | | -0.0020 | -0.0018 |
| 49 | Population Density x ic2 | | -0.0006 | -0.0009 |
| 50 | Population Density x ic3 | 1.2142e-07 | -1.1818e-04 | -5.3475e-05 |
| 51 | Population Density x ic4 | | 0.0037 | 0.0036 |
| 52 | Population Density x sc1 | 4.2873e-07 | -1.3142e-04 | -2.7350e-04 |
| 53 | Population Density x sc2 | | 0.0014 | 0.0011 |
| 54 | Population Density x sc3 | | -0.0007 | -2.7039e-05 |
| 55 | Population Density x sc4 | | -0.0021 | -0.0021 |
| 56 | Distinct SKUs/Basket | -0.0056 | | |
| 57 | Distinct Categories/Basket | 0.0019 | -4.4418 | -3.9342 |
| 58 | Number of Items/Basket | -0.0973 | | |
| 59 | $\hat{P}^1$ | | -710.7093 | -374.8441 |
| 60 | $\hat{P}^2$ | | 5045.1798 | 4173.3322 |
| 61 | $\hat{P}^3$ | | -4612.2068 | -4107.2023 |
| 62 | $\hat{P}^4$ | | -3929.0021 | -3377.1737 |
| 63 | $\hat{P}^5$ | | 231.1697 | 327.4122 |
| 64 | $\hat{P}^6$ | | 4139.4634 | 3558.4664 |
| 65 | Prop. of sales in LP category (pre LP) | 5.7896 | | |
| 66 | Received promotional email | 0.1106 | | |
| 67 | Received promotional email for LP | 1.7971 | | |
| 68 | Days until purchase in LP category | -0.0301 | | |
| 69 | Days until purchase in LP discount category | | | |

## 4.4 Visual Representation of Results

Even after variable reduction via LASSO, it is still difficult to concisely describe how the competitive structure influences estimated changes in spend due to the codependency between variables. We therefore visually illustrate the LASSO's results on how the competitive structure influences LP effectiveness. For brevity, we limit our discussion to two heatmaps that show the predicted change in spending for hypothetical customers. Additional heatmaps are provided in the appendix. These maps show how there is large heterogeneity across different customers with different competitive structures in the potential gains from having them join the LP. However, since these maps are mostly stylized representations of the actual data we focus on comparing the magnitudes across maps rather than specific prediction levels.

Figure 7 plots all four competitors, each set an equal distance from the focal store. The color on the heatmap reflects the estimated change in spend from a customer in that position if they were to join the focal store's loyalty program, with shades closer to white representing greater increases. This map shows that most of the gains are from customers near the big-box generalist, and less so from the small-box specialist. This map highlights that the value of a potential customer is heavily dependent on the extent to which the focal store can steal business from the competition. This also represents visually how a targeting strategy might be employed using complex representations of spatial relationships constructed from readily available data, even in the absence of past sales data on customers.

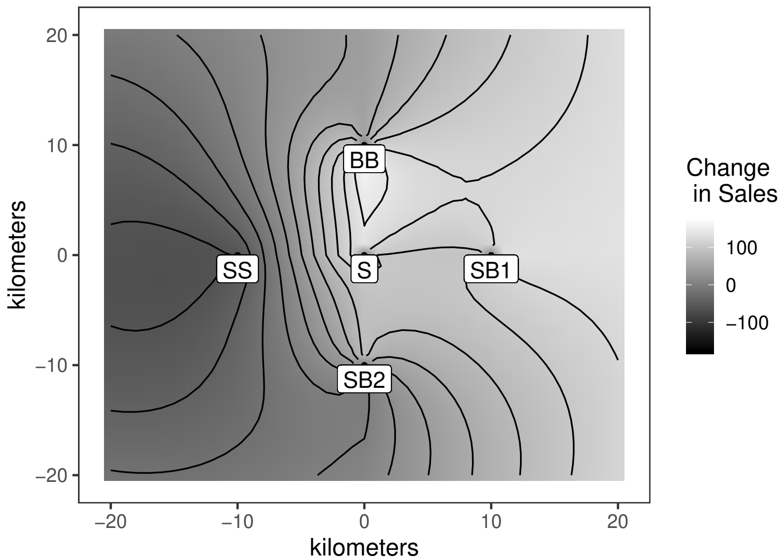



Figure 7: All Competitors

Figure 8 illustrates the predicted effects from an actual competitive structure randomly drawn from the data. Customers located on the far side of the second small-box specialist (SB2), and the big-box generalist (BB), have the lowest predicted change in spend. These customers are relatively far away from the focal store with a competitor directly in their path of travel, resulting in limited predicted gains. On the other hand, the other two competitors (SB1 and SS) are relatively close and the line of travel is not as critical with these competitors. The customers that show the most promise are those that are relatively close to the focal store but in the direction of the second small box generalist (SB2) and the big-box generalist (BB).

We also included two customer locations (I1 and I2) to illustrate more specifically the impacts of the competitive structure. The first customer, I1, is roughly halfway between the focal store and the second small box competitor (SB2). This customer has relatively high predicted change in spend in part, we suspect, of the potential to consolidate purchases. In contrast, the second customer, I2, is located right near two other competitors. Because this customer is so close to the competition, relative to the focal store, the expected change in sales is slightly dampened.

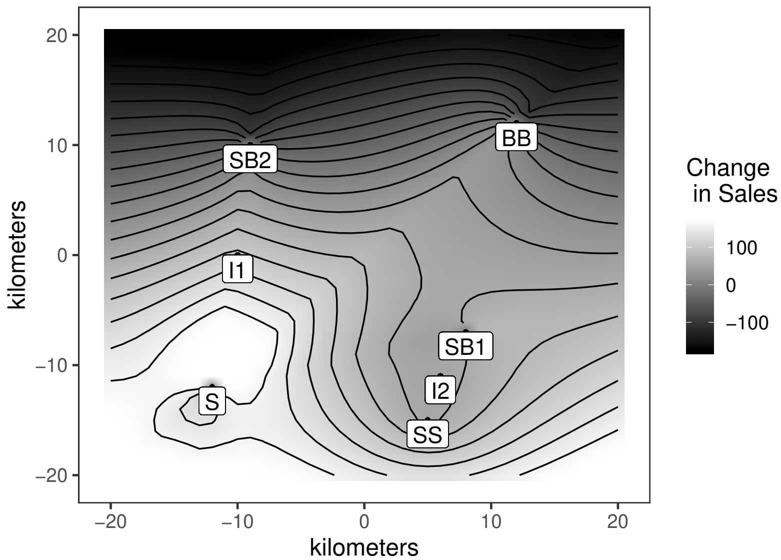



Figure 8: Randomly Selected Competitive Structure

The heatmaps presented convey two key advantages of this analysis. First, the geographic influence of a competitor is relatively complex. It makes clear how simple radii surrounding either the competitor or the focal store would not sufficiently capture the influence of the competitive structure on customer behavior. Second, these complex relationships are not fixed and can vary by the type of competitor in an area. This is an important consideration for the firm in the formation of targeting strategies: the combined relationship between the location of each competitor, customer, and focal store strongly influences the customer's interaction with the focal firm.

## 4.5 Competitive Segmentation: Out of Sample Validation

In this section we use the estimated model to demonstrate the value of augmenting traditional sales data with readily available spatial data on competitors. We use the LASSO results to form a segmentation strategy designed around predicting how customer spending changes upon joining the LP. Our aim is to motivate firms to augment their detailed customer transaction data with information on competitor locations to identify customers who are likely to be most profitable upon joining a LP. By identifying which segments are more likely to increase spending upon joining the LP the firm can focus its limited marketing budget on promoting the LP to those customers and avoid promoting it to those who are predicted to not alter spending patterns and simply take discounted prices. More generally, our results suggest that firms should consider segmenting on both horizontal quality and vertical quality to identify their most valuable customers.

To evaluate the potential increase in profits from this type of targeting, we use a validation or holdout sample of customers who were not used in estimation. For simplicity, we segment customers at the store level, rather than the firm level, since the competitive structure varies across each customer-store interaction. Our primary dependent variable of interest is the expected change in monthly spending.

Importantly, we only predict the estimated change in spend from customers who actually joined the LP. We then measure the success of each segmentation approach by calculating the actual change in spending from those customers whose predicted incremental profits from joining the LP is positive. In this way, we can evaluate each segmentation strategy by focusing on different subsets of customers who actually joined, acknowledging the fact that there may be self-selection into the LP. It also means we do not need to use the model to forecast how spending would change among customers we do not observe joining the LP, which may introduce noise into this exercise.

**Rule-of-Thumb Segmentation Strategies:**

The performance of numerous different potential segmentation rules are shown in Table 11. We first describe a series of simple segmentation criteria that are not based on the full competitive structure but instead rely on simple location or spending based rules-of-thumb. In each case we compare the performance of these rules by evaluating the total change in spending of all customers who would be identified by the rule.

First, a mass marketing strategy would be to simply avoid segmentation and treat the customers as a homogeneous group. In aggregate, customers show a slightly positive change in monthly spending after joining. This is consistent with the descriptive result that spending only mildly increases on average after customers join the LP.

We also consider a simple strategy based on vertical quality, consisting of focusing on only those that exhibit relatively high spending amounts (prior to joining the LP) at a few reasonable cutoff points (based on the empirical distribution of monthly spend at a single store location). As noted in Section 3, this is close to the actual way in which the focal firm conducts targeting. We consider three levels as a cutoff: \$80 (the median), \$100, and \$200 (about the 75th quantile). In each of these the change in monthly spend actually declines substantially. Segmentation in this manner would therefore be a highly unprofitable strategy.

Next we consider a naive location-based strategy that selects customers within a few kilometers of the focal store. If we limit attention to only those customers within 5 kilometers of the target store we get a small positive boost, which is only slightly stronger than a mass marketing approach. A second type of naive location-based segmentation would be to emphasize horizontal variables and try to take advantage of the business stealing aspect of loyalty programs by focusing on all customers located near a competitor. We test this strategy and find that customers within 5 km of a competitor exhibit a substantial decrease in profits.

**Model-Based Spatial Segmentation Strategies:**

Next, we compare several series of models distinguished by which data are used as inputs as well as how they are estimated. Each model is estimated on the training observations and validated on the holdout set. Note that we expect a higher degree of noise between the train and test set than in other applications, particularly LASSO applications, because we are not simply estimating change in spend but estimating a predictive model of change in spend and then grouping all customers with a predicted spend greater than zero. This comparison is more suited to the actual firm application of segmentation compared to a simple forecast of change in spend.

We start by focusing on different methods for including spatial data to derive some insight into how local competitive structure effects LP performance. First we test a simple spatial model that takes the standard approach to spatial data of only considering distances between the focal store and customer and each of the four competitors. We see that the out-of-sample performance is actually worse than a mass marketing approach. In other words, while horizontal data is valuable in principle a simplistic approach to incorporating it that misses the nuances of competitive structure is not necessarily effective.

The second and third models extend the simple model with additional spatial metrics. First we include the count of the number of each of the four competitors within a 10 kilometer (6 miles) radius, approximately the average driving distance. Then, we augment the original spatial model with the apex angle, or the angle formed between the customer, focal store, and a given competitor with the customer at the apex. The results show that also knowing the number of nearby competitors is much more valuable than simply knowing distance between the customer and each store. In addition, knowing the angle between customer and competitors improves substantially over simply knowing distance. In terms of comparing these two data types, they perform roughly similarly as one another.

Next, to provide another comparison, we estimate three spatial models in the spirit of classic gravity models. These models date back to Reilly (1931) who coined the "law of retail gravitation" and are extended to a more flexible model by Huff (1964). These gravity models are designed to predict store choice while we have richer data on spending, but only at the focal store. We attempt to preserve the theoretical insight of these models but adjust the specification to instead predict the probability of a positive change in spend with $p_{k=1} = \frac{w_1/d_1^{\alpha}}{\sum_{k=1}^{5} w_k/d_k^{\alpha}}$, where $d$ represents the distance between the customer and each store type $k$ (the focal store and the four competitors) and $w$ are the weights or "attractiveness" of each store.

We consider several Huff-style model specifications. In the first, we set the store attractiveness or mass to be the square footage of each competitor type and in the second specification we treat each store attractiveness as a parameter to estimate. In both of these cases, the decay parameter $\alpha$ is estimated using maximum likelihood. In our third specification we allow $\alpha_k$ to differ over store type. For segmentation, we take the models predictions and group the customers based on the percentile of positive spend in the training data.[21]

We find that the gravity models perform quite well overall at identifying profitable customer groups. They outperform the mass marketing approach and the rule of thumb strategies. They also do better out-of-sample than the models using simple sales data or distance data alone, suggesting that those models are prone to overfitting in a way the more parsimonious gravity models are not. However, performance is similar to the augmented simple spatial models. In the next subsection we show that our preferred specification performs about three times as as well as the Huff-style gravity models, however.

Finally, we compare each of these spatial models to a model that considers only vertical quality or historical sales information: overall sales, category level sales, and category level sales that specifically qualify for the LP. This model also under performs a mass segmentation approach and results in negative profits, suggesting that these data alone have poor predictive power and are prone to overfitting.

**LASSO Model-Based Segmentation Strategies:**

Finally, we estimate a series of models using our preferred LASSO approach.

[21]We attempted numerous variations on the Huff model given that it is not a direct translation of the original method, and kept only those that performed best at predicting change in spending. More details are available upon request.

The first model uses the LASSO approach but ignores competitor spatial information. We estimate the model using traditional RFM metrics alone. This model outperforms all others thus far both in-sample and out-of-sample, attesting to the strength of using the LASSO approach as a baseline for improving predictive power.

Next, we show results for our preferred spatial LASSO. Here we augment the LASSO RFM model with the spatial information of the competition. It outperforms all other segmentation strategies in the validation sample, where the customers identified exhibit an actual change in sales nearly ten times higher than those in a mass marketing targeting scheme. The improvements compared to the RFM LASSO model are also substantial: about a 19% increase in the out-of-sample change in sales and 29% change in non-qualified sales. In short, the spatial LASSO does the best job of predicting which LP joiners end up being most valuable.

Finally, we compare the performance of an RFM approach to a spatial approach when estimated and validated separately in competitive and isolated markets.[22] The spatial model outperforms the RFM only model in all comparisons where the market type (isolated or competitive) or data subset (train or test) is held constant.

The comparisons across model inputs and market types are informative. The RFM only model performs nearly as well as the spatial model at identifying more valuable customers in isolated markets but substantially worse in competitive markets. The spatial model, by contrast, performs similarly in competitive markets as in isolated markets. This shows how the data on vertical characteristics adds predictive value mostly in settings where customers are not vulnerable to competition. The spatial data adds value regardless of market type.

These results are broadly informative on how 2nd degree price discrimination (quantity discounts or behavior-based pricing) and 3rd degree price discrimination (targeted price discounts) interact in practice. Properly designed quantity discounts should be profitable by increasing spending by high value customers. Therefore, in isolated markets it is effective to target these discounts using a measure of vertical quality constructed from past sales data. The quantity discounts work by letting customers self-select into the discounts based on their level of spending and marginal demand. In markets with close competitors, targeting can be done based on the horizontal vulnerability dimension and within the targeted group customers will self-select into the discounts based on vertical characteristics. It is better then to target customers who might shift their purchases away

[22] These market types are defined in section 3.

from competitors than to target the highest value customers alone, since they may not change their spending at all in response to the LP discounts.

| | Avg. Δ Sales | | Avg. Δ Non-Qual. Sales | |
|---|---|---|---|---|
| Targeting Strategy | Train | Test | Train | Test |
| Mass Marketing | -$18.69 | $6.91 | -$24.32 | $2.59 |
| Top spender (>$80) | -$83.07 | -$38.10 | -$88.18 | -$33.29 |
| Top spender (>$100) | -$95.38 | -$45.16 | -$99.92 | -$40.19 |
| Top spender (>$200) | -$181.13 | -$103.67 | -$178.76 | -$95.73 |
| Within 5 kilometers of focal store | $46.60 | $12.60 | $38.86 | $13.23 |
| Within 5 kilometers of any competitor | -$47.57 | -$54.54 | -$53.00 | -$41.93 |
| Simple spatial | $40.72 | $3.97 | $31.30 | $3.69 |
| Simple spatial w/ competitors in 10km radius | $38.85 | $31.91 | $33.09 | $27.77 |
| Simple spatial w/ apex angle | $40.09 | $26.73 | $33.32 | $20.27 |
| Gravity fixed $w$ ($\hat{\alpha}$ = .227, .232) | $13.76 | $20.16 | $13.11 | $22.34 |
| Gravity estimated $w$ ($\hat{\alpha}$ = .039, .041) | $11.05 | $24.90 | -$4.28 | $30.47 |
| Gravity $\alpha_k$ | $17.93 | $19.43 | $17.80 | $13.90 |
| Historical sales only | $29.06 | -$23.06 | $36.52 | -$12.46 |
| LASSO RFM only | $79.87 | $57.75 | $68.60 | $42.25 |
| LASSO RFM & Spatial | $84.09 | $68.44 | $77.40 | $54.49 |
| LASSO RFM only (Competitive) | $59.78 | $31.77 | $57.83 | $32.36 |
| LASSO RFM only (Isolated) | $87.04 | $65.70 | $72.44 | $45.28 |
| LASSO RFM & Spatial (Competitive) | $90.35 | $69.39 | $86.69 | $66.64 |
| LASSO RFM & Spatial (Isolated) | $82.28 | $68.09 | $74.72 | $49.97 |

Table 11: Targeting Validation Performance

In Table 12 we investigate the influence of our control function approach on targeting performance. We see that both for the full model and in the RFM-only model the selection correction generates a significant increase in out-of-sample predictive power. We consider this an important comparison since out-of-sample prediction is the ultimate goal of the model. We also note that the relative performance of the models with and without spatial data is made stronger by the use of selection correction.

| Targeting Strategy | Avg. Δ Sales Train | Avg. Δ Sales Test | Avg. Δ Non-Qual. Sales Train | Avg. Δ Non-Qual. Sales Test |
|---|---|---|---|---|
| **LASSO RFM only** | | | | |
| No selection correction | \$64.11 | \$45.16 | \$55.37 | \$33.93 |
| With selection correction | \$79.87 | \$57.75 | \$68.60 | \$42.25 |
| **LASSO RFM & Spatial** | | | | |
| No selection correction | \$77.04 | \$47.06 | \$68.86 | \$31.66 |
| With selection correction | \$84.09 | \$68.44 | \$77.40 | \$54.49 |

Table 12: LASSO Targeting with and without Selection Correction

One argument might be that since marketing is done via email and is relatively inexpensive, why not simply send marketing campaigns to everyone and not bother with segmenting the customers? The key here is that in some cases it might be preferable to limit LP enrollment to certain segments. While the firm cannot prevent interested customers from joining its loyalty program, it can choose which customers to invest in after they join (e.g., special offers). Identifying the right type of customer for this type of promotion can substantially increase its profitability. While this has generally been known, we show two distinct contributions. First, location data adds substantially more predictive power to a segmentation strategy than historical sales patterns. Second, simply incorporating traditional spatial metrics like distance between customer and store does not capture the complexity of competitive structure. Instead, we show how firms could take advantage of rich spatial data on both customers and competitors locations using a more sophisticated but still straightforward approach.

**Comparison of Sales Predictions by Customer Type:**

To better understand the contribution of horizontal variables in prediction exercise, we show how the different segmentation strategies would select customers who differ on vertical and horizontal metrics.

Table 13 presents the simplest spatial metrics (distances between the customer, focal store, and each competitor) and monthly spending of the customers targeted in each strategy. Since there is considerable overlap between the customers across these strategies, we parse out the observations that were predicted to increase sales in the spatial LASSO but not by the alternative segmentation strategies. In the table we label these as "missed" opportunities.

We see that customers predicted to increase spending by the model using traditional RFM metrics alone have lower pre-join sales per month relative to customers selected by the spatial LASSO and are positioned significantly further from the focal store. This finding is made evident when looking at the missed opportunities: these customers are much closer to the focal store and the sales per month is much higher. The other distance metrics show that these customers and store locations also tend to be closer to the competition. This emphasizes one of the key insights of the spatial model: on average, regardless of location, customers with higher than average sales per month are unlikely to show large increases in spend after joining the LP (likely because they are already dedicating most spend to the focal store). However, we posit that customers located relatively closer to both the focal store and the competition are more likely to consolidate purchases away from the competition and will exhibit higher than average changes in spend upon joining.

We next evaluate the customers from the rule of thumb segmentation strategy that focuses on customers close to the focal store. This illustrates a similar issue as the previous strategy. On average, customers close to the focal store are likely to have access to a competitor as well, so at first glance these customers appear to be promising prospects for targeting. However, they may be *too* close to the focal store, in the sense that they already dedicate a substantial amount of spend to the focal store (as evidenced by the relative high sales per month). The spatial LASSO recognizes that customers relatively further away may prove to be promising, so long as the have relatively similar access to competitors.

Finally, we consider the second Huff-style gravity model with estimated weights $w$, since, of the gravity models, this variation best segmented customers. Under this model, valuable customers are excluded simply because they are relatively far away from the focal store. However, for these customers all four competitors are closer to the focal store, compared with customers identified by the Huff model. We again find that the most valuable customers, in terms of expected change in spend, are those with the right balance distance between them and the focal store locations and the competition. This highlights a potential weakness of a traditional gravity model, which does not explicitly consider distances between the store locations, only the distances to the customer.

This analysis shows how the spatial LASSO model can capture more subtle drivers of customer behavior that are typically missed when looking at transaction data alone (e.g., RFM variables) or relatively simple spatial components. We find that there are substantial gains in segmentation actions by accounting for more complex spatial relationships among the customer, focal store, and relevant competitors.

| | Out-of-Sample Performance | | | | | | |
|---|---|---|---|---|---|---|---|
| | | LASSO (RFM only) | | w/i 5km of Focal Store | | Gravity (est. $w$) | |
| | LASSO | Targeted | Missed | Targeted | Missed | Targeted | Missed |
| Δ Monthly Sales | \$68.44 | \$56.94 | \$40.71 | \$12.60 | \$61.64 | \$24.90 | \$69.71 |
| Δ Qualified Sales | \$13.95 | \$15.77 | \$-2.07 | \$-0.63 | \$21.63 | \$0.15 | \$24.12 |
| Δ Non-Qualified Sales | \$54.49 | \$41.17 | \$42.78 | \$13.23 | \$40.01 | \$24.75 | \$45.58 |
| Sales/Month (pre join) | \$185.14 | \$118.56 | \$427.23 | \$424.25 | \$108.60 | \$385.13 | \$85.63 |
| Sales/Month (pre join) SD | \$343.87 | \$198.79 | \$550.03 | \$534.57 | \$188.89 | \$482.54 | \$127.87 |
| Sales/Month (pre join) 25th | \$21.43 | \$20.11 | \$132.95 | \$107.78 | \$15.63 | \$91.43 | \$15.48 |
| Sales/Month (pre join) 75th | \$155.24 | \$115.28 | \$409.69 | \$507.74 | \$107.06 | \$473.77 | \$85.62 |
| Sales/Month (post join) | \$252.41 | \$181.35 | \$467.94 | \$430.11 | \$171.67 | \$398.32 | \$166.42 |
| si: Store-Customer (km) | 9.29 | 13.35 | 5.75 | 3.00 | 12.67 | 5.20 | 14.00 |
| ic1: Customer-BB (km) | 14.17 | 14.73 | 15.36 | 13.00 | 14.80 | 14.28 | 13.31 |
| ic2: Customer-SB1 (km) | 14.00 | 16.76 | 11.71 | 15.55 | 14.09 | 14.44 | 15.38 |
| ic3: Customer-SB2 (km) | 16.69 | 16.81 | 19.89 | 14.98 | 18.37 | 17.14 | 16.96 |
| ic4: Customer-SS (km) | 14.94 | 16.26 | 14.23 | 13.62 | 15.69 | 14.69 | 14.98 |
| sc1: Store-BB (km) | 14.34 | 14.35 | 15.61 | 12.78 | 15.12 | 15.01 | 12.76 |
| sc2: Store-SB2 (km) | 14.42 | 16.29 | 12.97 | 15.47 | 14.71 | 14.34 | 15.91 |
| sc3: Store-SB1 (km) | 16.25 | 17.57 | 19.50 | 14.92 | 17.67 | 18.02 | 15.17 |
| sc4: Store-SS (km) | 14.52 | 14.82 | 14.18 | 13.55 | 15.00 | 15.15 | 13.46 |

Table 13: Targeted Customers

## 4.6 Spatial Variable Contribution

In the previous section, we emphasized the value of integrating spatial or horizontal information into the segmentation process. In this section we provide an additional measure of the contribution of spatial variables to the model fit relative to other types of variables. Specifically, we highlight that our spatial metrics contribute substantially more towards $R^2$ relative to traditional vertical variables.

Table 14 presents the model fit using our two dependent variables of interest, where each row represents a single subset of the available variables. The first column displays the pseudo-$R^2$ from a logistic regression and the second column the standard $R^2$ from a linear regression. The variables are separated into one of four labels: our proposed spatial metrics, RFM (trip frequency and sales information), basket composition (number of items per basket, distinct items, distinct categories, etc.), and finally the marketing indicators.

The spatial metrics explain considerably more than any of the alternative variable types. The sum of the $R^2$ values across the partial models is similar to the fit from the aggregate model in the final row. This suggests that the variable types tend to be relatively orthogonal to each other. In other words, the spatial information appears to be adding a non-trivial amount of unique information to the model that would otherwise be absorbed into the residuals.

A natural rebuttal to this table is that the results are driven simply by the sheer number of spatial variables relative to the other variable types. This is only partially true: recall that all of the spatial metrics are derived solely from the latitude and longitude of the focal store, competition, and the customer. From a managerial perspective it is very easy to recreate the diverse set of spatial metrics once these few location points are obtained. In this sense, it is more the variety among the spatial variables we designed, rather than simply the number of variables, that contributes a substantial amount of information to each model.

| **Variable Type** | **Pr(Join)** | **$\Delta$ Sales\|Join** |
|---|---:|---:|
| Spatial | 3.4% | 14.9% |
| RFM | 0.6% | 9.8% |
| Basket composition | 1.3% | 3.7% |
| Marketing | 0.1% | 0.0% |
| All variable types | 5.3% | 27.0% |

Table 14: Pseudo-$R^2$ and $R^2$ by variable type

# 5 Discussion and Conclusion

Both academics and managers have been interested in understanding the factors contributing to a successful loyalty program. We use a large new dataset on customer spending behavior at a major U.S. retailer to develop a better understanding of the role of spatial data and competitive structure in LP performance and in the effectiveness of price discrimination strategies more broadly. We argue that a significant but largely overlooked driver of a loyalty program's success is the joint spatial relationship among the focal firm's stores, the competition, and the customer, or what we collectively label the competitive structure.

We first document large changes in behavior after joining the LP, with a wide degree of heterogeneity across customers. Notable among this customer heterogeneity are customer segments we describe as consolidators or as upgraders. This heterogeneity suggests that targeting the LP with promotions could be effective in dealing with the well-known issue that in many cases firms believe that LPs fail to increase profits.

Second, we find strong model-free evidence that access to competitor stores is the strongest predictor of whether a customer will increase spending upon joining the LP. This suggests that LPs using the common form of quantity discounts work via business stealing rather than overall demand expansion. Third, in informal comparisons we show that past spending behavior is not very predictive of how spending will change when a customer joins the LP. These data are less predictive than spatial data and in particular the more complex representations of spatial relationships.

We then develop formal models of how spending changes. This provides several important results. First, segmentation for providing quantity discounts (or performing second degree price discrimination) naively to the highest spenders would be very unprofitable. Naively segmenting customers based on simple spatial rules would not be very effective either, performing roughly the same as not treating all customers the same. Next, using LASSO to select among many possible relationships in the estimation data helps predict changes in spending substantially better than simple models. Then, comparing performance of LASSO models relying on vertical variables as inputs to using horizontal variables shows a substantial improvement associated with the readily available spatial data.

These results suggest a segmentation strategy for firms interested in identifying their most profitable customers. They are also informative regarding how and when second degree price discrimination or one type of behavior-based pricing are likely to be successful. As Shin and Sudhir (2010) note, it can profitable to use this type of pricing strategy if there exists substantial heterogeneity in customer value and customer preferences are stochastic. Our results confirm this insight and operationalize it by defining a customer's vertical quality using past spending data and their horizontal quality using spatial relationships to competitors and the focal store. Quantity discounts based on vertical quality can be effective in relatively isolated markets but in competitive settings targeting on horizontal variables is much more effective.

Still, our analysis is not without limitations. Due to the nature of the data we cannot make very strong statements on the extent of business stealing that may be occurring. In the absence of data from multiple competitors or direct responses from the customers themselves, we can only infer the degree of business stealing based on changes in shopping behavior as a function of competitive structure. We hope that our research provides the foundation for future work that integrates the complex spatial information into models of customer behavior, especially in their relationship to making more strategic firm decisions.

## References


Ahn, H. and Powell, J. L. (1993). Semiparametric estimation of censored selection models with a nonparametric selection mechanism. *Journal of Econometrics*, 58(1-2):3–29.

Ascarza, E. (2018). Retention futility: Targeting high-risk customers might be ineffective. *Journal of Marketing Research*, 55(1):80–98.

Belloni, A., Chernozhukov, V., et al. (2013). Least squares after model selection in high-dimensional sparse models. *Bernoulli*, 19(2):521–547.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, 81(2):608–650.

Besanko, D., Dube, J.-P., and Gupta, S. (2003). Competitive price discrimination strategies in a vertical channel using aggregate retail data. *Marketing Science*, 49(9):1121–1138.

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

Blattberg, R. C. (2008). Byung-do kim, and scott a. neslin (2008), database marketing: Analyzing and managing customers. *International Series in Quantitative Marketing*.

Chen, Y., Li, X., and Sun, M. (2016). Competitive mobile geo targeting. *Marketing Science*, 36(5):645–812.

Chen, Y., Narasimhan, C., and Zhang, Z. J. (2001). Individual marketing with imperfect targetability. *Marketing Science*, 20(1):1–98.

Chen, Y. and Steckel, J. H. (2012). Modeling credit card share of wallet: Solving the incomplete information problem. *Journal of Marketing Research*, 49(5):655–669.

Cook, J. A. and Siddiqui, S. (2020). Random forests and selected samples. *Bulletin of Economic Research*, 72(3):272–287.

DeWulf, K., Odekerken-Schröder, G., and Iacobucci, D. (2001). Investments in consumer relationships: A cross-country and cross-industry exploration. *Journal of Marketing*, 65(4):33–50.

Dreze, X. and Nunes, J. C. (2008). Feeling superior: The impact of loyalty program structure on consumers' perceptions of status. *Journal of Consumer Research*, 35(6):890–905.

Du, R. Y., Kamakura, W. A., and Mela, C. F. (2007). Size and share of customer wallet. *Journal of Marketing*, 71(2):94–113.

Dube, J.-P., Fang, Z., Fong, N., and Luo, X. (2017). Competitive price targeting with smartphone coupons. *Marketing Science*, 36(6):813–1017.

Ellickson, P. and Misra, S. (2012). Enriching Interactions: Incorporating Revenue and Cost Data into Static Discrete Games. *Quantitative Marketing and Economics*, 10:1–26.

Fong, N. M., Fang, Z., and Luo, X. (2015). Geo-conquesting: Competitive locational targeting of mobile promotions. *Journal of Marketing Research*, 52(5):726–735.

Frank, L. E. and Friedman, J. H. (1993). A statistical view of some chemometrics regression tools. *Technometrics*, 35(2):109–135.

Ghose, A. and Yang, S. (2009). An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management science*, 55(10):1605–1622.

Hartmann, W. and Viard, B. (2008). Do frequency reward programs create switching costs? a dynamic structural analysis of demand in a reward program. *Quantitative Marketing and Economics*, 6(2):109–137.

Heckman, J. J. and Robb, R. (1986). Alternative methods for solving the problem of selection bias in evaluating the impact of treatments on outcomes. In *Drawing inferences from self-selected samples*, pages 63–107. Springer.

Heckman, J. J. and Robb Jr, R. (1985). Alternative methods for evaluating the impact of interventions: An overview. *Journal of econometrics*, 30(1-2):239–267.

Huff, D. L. (1964). Defining and estimating a trading area. *The Journal of Marketing*, pages 34–38.

Jain, A., Misra, S., and Rudi, N. (2020). The effect of sales assistance on purchase decisions: An analysis using retail video data. *Quantitative Marketing and Economics*, pages 1–31.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to statistical learning*, volume 112. Springer.

Leenheer, J., Van Heerde, H. J., Bijmolt, T. H., and Smidts, A. (2007). Do loyalty programs really enhance behavioral loyalty? an empirical analysis accounting for self-selecting members. *International Journal of Research in Marketing*, 24(1):31–47.

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

Li, Y., Gordon, B., and Netzer, O. (2018). An empirical study of national vs. local pricing by chain stores under competition. *Marketing Science*, 37(5):812–837.

Liu, Y. and Yang, R. (2009). Competing loyalty programs: Impact of market saturation, market share, and category expandability. *Journal of Marketing*, 73(1):93–108.

Lockhart, R., Taylor, J., Tibshirani, R. J., and Tibshirani, R. (2014). A significance test for the lasso. *Annals of statistics*, 42(2):413.

Luo, X., Andrews, M., Fang, Z., and Phang, C. W. (2014). Mobile targeting. 60(7):1738â56.

McCall, M. and Voorhees, C. (2010). The drivers of loyalty program success: An organizing framework and research agenda. *Cornell Hospitality Quarterly*, 51(1):35–52.

Meyer-Waarden, L. and Benavent, C. (2009). Grocery retail loyalty program effects: self-selection or purchase behavior change? *Journal of the Academy of Marketing Science*, 37(3):345–358.

Nambiar, M., Simchi-Levi, D., and Wang, H. (2019). Dynamic learning and pricing with model misspecification. *Management Science*, 65(11):4980–5000.

Phillips, R., Şimşek, A. S., and Van Ryzin, G. (2015). The effectiveness of field price discretion: Empirical evidence from auto lending. *Management Science*, 61(8):1741–1759.

Reilly, W. (1931). The law of retail gravitation. wj reilly. *Inc., New York*.

Roehm, M. L., Pullins, E. B., and Roehm Jr, H. A. (2002). Designing loyalty-building programs for packaged goods brands. *Journal of Marketing Research*, 39(2):202–213.

Rossi, P., McCulloch, R., and Allenby, G. (1996). The value of purchase history data in target marketing. *Marketing Science*, 15(4):301–394.

Shaffer, G. and Zhang, Z. J. (1995). Competitive coupon targeting. 14(4):343–466.

Shin, J. and Sudhir, K. (2010). A customer management dilemma: When is it profitable to reward one's own customers? *Marketing Science*, 29(4):671–689.

Shugan, S. (2005). Brand loyalty programs: Are they shams? *Marketing Science*, 24(2):185–193.

Stourm, V., Bradlow, E., and Fader, P. (2017). Market positioning using cross-reward effects in a coalition loyalty program. *HEC Paris Research Paper No. MKG-2017-1242*.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288.

Verhoef, P. (2003). Understanding the effect of customer relationship management efforts on customer retention and customer share development. *Journal of Marketing*, 67(4):30–45.

Wachtel, S. and Otter, T. (2013). Successive sample selection and its relevance for management decisions. *Marketing Science*, 32(1):170–185.

Wirtz, J., Mattila, A. S., and Oo Lwin, M. (2007). How effective are loyalty reward programs in driving share of wallet? *Journal of service Research*, 9(4):327–334.

Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Xu, X., Li, X., and Zhang, J. (2020). Regularization methods for high-dimensional sparse control function models. *Journal of Statistical Planning and Inference*, 206:111–126.

# Appendix 1: Distance Metrics

In the table below, $x$ denotes the competitor type. In our empirical application this is either the big box generalist (BB), the small box generalists (SB1 or SB2), or the small box specialist (SS).

| **Distance Metric** | **Description** |
|---|---|
| si | Distance between the focal store s and individual i (in km) |
| ic$x$ | Distance between the individual and competitor $x$ (in km) |
| sc$x$ | Distance between the focal store and competitor $x$ (in km) |
| s$x$ | Sparsity metric for competitor $x$, defined as sc$x$/(ic$x$ + si) |
| nAll_$x$ | Total number of $x$ competitors within the region |
| n$d$_$x$ | Total number of $x$ competitors within $d$ km of the individual |
| cInt$x$ | Angle formed between focal store and competitor with individual at apex<br>(0 means s and $x$ are in same direction relative to c<br>180 means s in opposite direction of $x$) |
| icbar_$x$ | Average km between the individual and each individual competitor $x$<br>(if only one competitor of type $x$ in an area, this equals ic$x$) |
| ccbar_$x$ | Average km distance between each competitor and the competitive center<br>(if only one competitor of type $x$ in an area, this equals 0) |

Table A.1: Distance Metric Descriptions

## Appendix 2: Supplementary Marketing Analysis

Here we provide additional details on the firm's marketing activity and its relationship to the location of the customers and the LP. There are two primary issues to consider: If marketing interventions change when customers join the loyalty program, then the change in behavior that we observe post-join could be at least partially caused by these marketing actions and not the LP itself. So we would overstate the effect of the LP membership on behavior. A second important concern is that if these marketing actions are targeted to specific customers based on their characteristics (like location), then the differences across customers in how behavior changes after joining the LP are potentially related to marketing actions and not the LP. This would be a concern for our segmentation strategy.

While this analysis is informative, it is somewhat tangential to the primary goal of the paper hence its placement here.

**How is marketing targeted?** First, we look at the average spending and trip frequency for customers who are sent marketing emails vs those who are not. For those that receive marketing, we use the level of spending and trip frequency prior to receiving their first marketing email to avoid reverse causation. In the table below, we see a clear relationship between the type of marketing received (if any) and the customers' value.

| Marketing Group | Monthly Spend (Pre-LP if Joined) | Trips/Month |
|---|---|---|
| No Emails | \$373 | 3.9 |
| Regular Emails Only | \$560 | 4.7 |
| Join LP Emails | \$695 | 5.7 |

Table A.2: Monthly Spend by Marketing Group

For those who receive marketing emails, we also look at the correlation between the *number* of emails they receive and their level of spending and frequency of visit, conditional on receiving at least one email. We find that there is no relationship between the number of marketing emails received and monthly spend or trip frequency, this is true for both types of emails: generic emails and those specifically encouraging a join to the LP.

So there is clearly some targeting of marketing actions to customers with relatively high spending levels and trip frequencies. Next we look at the relationship between marketing and a customer's spatial relationship to the focal store and its competitors. We first calculated the correlations between all of the spatial variables and the marketing activity variables and found that no meaningful relationships, with a maximum coefficient of determination ($R^2$) of .002 across all variables both types of marketing emails. These correlations are shown in the figure below to illustrate the very small magnitudes of the correlations, especially in the traditional distance metrics (i.e., metrics we did not develop and therefore more likely to be used by the firm).

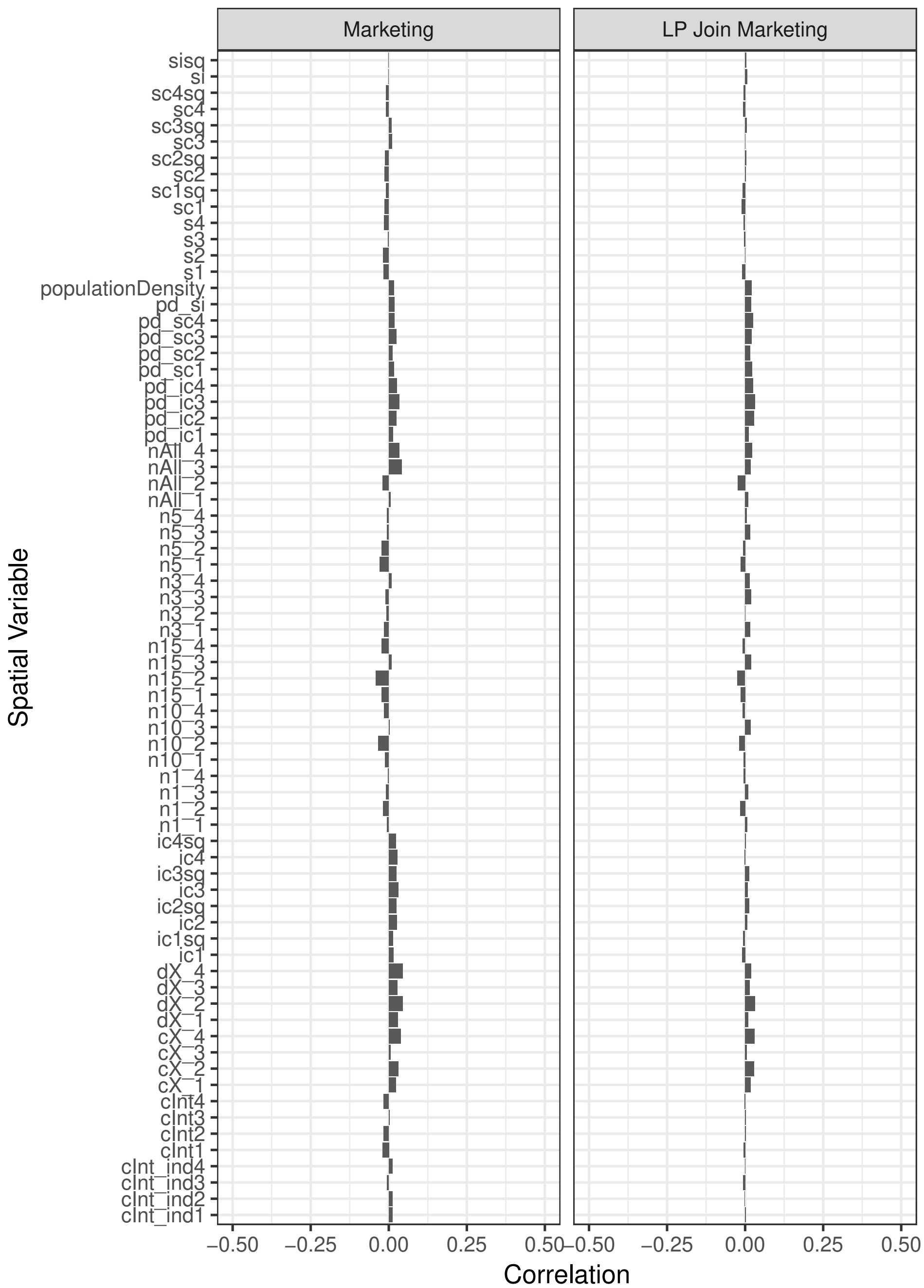



Figure A.1: Correlations between Spatial and Marketing Variables

In addition, we ran regressions on predicting marketing activity of both types of emails (generic and those encouraging an LP join) using basic spatial information (i.e., metrics that we did not develop and would thus not likely to be used by the firm). For both types of emails, we estimated the propensity to receive an email, and the email frequency among those that received at least one email. The results in the table below show that there is very limited evidence that the firm's marketing activity are related to spatial information in any systematic way.

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Email | Join LP Email | Email Freq.\|Email | Join LP Freq.\|Email |
| Constant | $-1.047^{***}$ | $-2.896^{***}$ | $8.502^{***}$ | $0.029^{**}$ |
| | (0.147) | (0.304) | (1.108) | (0.012) |
| si | 0.005 | 0.012 | 0.020 | 0.0003 |
| | (0.006) | (0.013) | (0.044) | (0.0005) |
| ic1 | $-0.006$ | $-0.011$ | $0.066^{*}$ | $-0.00003$ |
| | (0.005) | (0.011) | (0.038) | (0.0004) |
| ic2 | 0.005 | 0.005 | 0.040 | 0.0002 |
| | (0.004) | (0.007) | (0.027) | (0.0003) |
| ic3 | 0.003 | 0.003 | $-0.027$ | $-0.0004$ |
| | (0.003) | (0.006) | (0.023) | (0.0003) |
| ic4 | 0.002 | $-0.005$ | $-0.064$ | $-0.001$ |
| | (0.006) | (0.013) | (0.046) | (0.001) |
| sc1 | 0.009 | $-0.011$ | $-0.016$ | 0.00001 |
| | (0.007) | (0.016) | (0.056) | (0.001) |
| sc2 | $-0.017^{***}$ | $-0.009$ | 0.020 | 0.0003 |
| | (0.006) | (0.012) | (0.046) | (0.001) |
| sc3 | $-0.003$ | 0.001 | $-0.003$ | $0.001^{*}$ |
| | (0.004) | (0.009) | (0.034) | (0.0004) |
| sc4 | 0.002 | 0.004 | 0.016 | $-0.001$ |
| | (0.009) | (0.019) | (0.067) | (0.001) |
| Population Density | $0.00004^{***}$ | $0.00003^{*}$ | $0.0001^{*}$ | $-0.00000$ |
| | (0.00001) | (0.00002) | (0.0001) | (0.00000) |
| Observations | 9,604 | 9,604 | 2,750 | 2,750 |
| $R^2$ | | | 0.005 | 0.006 |
| Log Likelihood | $-5{,}730.304$ | $-1{,}919.136$ | | |
| *Note:* | | | | $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$ |

Table A.3: Email Activity Related to Spatial Information

These findings lead to our conclusion that the firm does not target on spatial relationships, only on spending levels.

**How does marketing change post-LP?** Next we explore how marketing changes after customers join the LP. To begin, we emphasize an important aspect of our data is that we observe customer spending patterns both before and after they join the LP. The firm tracks customers through their credit card spend, using matching techniques to combine data across multiple credit cards or purchase methods (e.g., same name and address on different credit cards). So it is not the case for this firm that they only have detailed data on its customers after they join the LP. This means that implicitly there is no reason to think that marketing should or would change after they join the LP, and because of the panel structure of the data this is also something that can be tested.

In general, the email frequency increases for all customers over time, regardless of join status. Further, the email frequency between joiners and non-joiners is nearly identical over time. This was illustrated in Figure 3 from Section 3, which shows that overall email frequency (regular and LP emails) between joiners and non-joiners moves in tandem over time, suggesting that the firm does not alter their marketing strategy for joiners. This was also confirmed in a regression of email frequency on average spending, trips, LP membership status, and weeks since the start of the data. Neither the LP status nor the average spending levels are significantly related to email frequency.

We ultimately find no evidence that the firm changes its marketing actions for LP-joiners after they become a member of the LP. Again, this is not surprising because the firm observes spending data on its customers even if they are not LP members and so the primary reason the firm would change its marketing (additional customer data) is not a factor here.

## Summary of Marketing Analysis

First, the firm seems to target marketing actions to higher spending customers. This is true for both regular marketing promotion emails and emails designed to encourage the customer to join the LP. Second, we cannot find any relationship between the firm's marketing actions and customer location relative either to its own stores or competitor stores. Third, we do not find evidence that marketing actions change after customers join the LP.

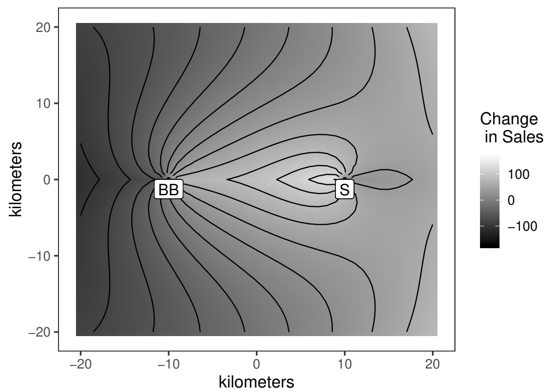
We conclude from this that the changes in spending we observe after customers join the LP are not a result of changes in marketing received, and in particular the strong relationship we find between a customer's local competitive structure and the amount they change spending after joining the LP cannot be explained by marketing actions.

| | *Dependent variable:* | | | | | |
|---|---|---|---|---|---|---|
| | y_j | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| (Intercept) | $-3.507^{***}$ (0.158) | $-3.470^{***}$ (0.157) | $-3.469^{***}$ (0.156) | $-3.463^{***}$ (0.156) | $-3.447^{***}$ (0.156) | $-3.448^{***}$ (0.156) |
| ic2 | 0.0004 (0.005) | 0.0002 (0.005) | 0.001 (0.005) | 0.001 (0.005) | 0.001 (0.005) | 0.001 (0.005) |
| ic3 | $0.005^{**}$ (0.002) | $0.005^{**}$ (0.002) | $0.005^{**}$ (0.002) | $0.005^{**}$ (0.002) | $0.005^{**}$ (0.002) | $0.005^{**}$ (0.002) |
| ic2sq | 0.0002 (0.0001) | 0.0002 (0.0001) | 0.0002 (0.0001) | 0.0002 (0.0001) | 0.0002 (0.0001) | 0.0002 (0.0001) |
| ic4sq | $-0.0001$ (0.0001) | $-0.0001^{*}$ (0.0001) | $-0.0002^{*}$ (0.0001) | $-0.0002^{*}$ (0.0001) | $-0.0002^{*}$ (0.0001) | $-0.0002^{*}$ (0.0001) |
| sc1 | 0.007 (0.005) | 0.007 (0.005) | 0.006 (0.005) | 0.006 (0.005) | 0.006 (0.005) | 0.006 (0.005) |
| sc1sq | 0.0001 (0.0001) | 0.0001 (0.0001) | 0.0001 (0.0001) | 0.0001 (0.0001) | 0.0001 (0.0001) | 0.0001 (0.0001) |
| sc3sq | $0.0001^{**}$ (0.00004) | $0.0001^{**}$ (0.00004) | $0.0001^{**}$ (0.00003) | $0.0001^{**}$ (0.00004) | $0.0001^{**}$ (0.00004) | $0.0001^{**}$ (0.00004) |
| sc4sq | $-0.0002^{**}$ (0.0001) | $-0.0002^{***}$ (0.0001) | $-0.0002^{**}$ (0.0001) | $-0.0002^{**}$ (0.0001) | $-0.0002^{**}$ (0.0001) | $-0.0002^{**}$ (0.0001) |
| s4 | $-0.465^{***}$ (0.117) | $-0.461^{***}$ (0.117) | $-0.464^{***}$ (0.116) | $-0.468^{***}$ (0.116) | $-0.473^{***}$ (0.116) | $-0.470^{***}$ (0.116) |
| nAll_1 | $0.020^{***}$ (0.003) | $0.020^{***}$ (0.003) | $0.020^{***}$ (0.003) | $0.020^{***}$ (0.003) | $0.020^{***}$ (0.003) | $0.020^{***}$ (0.003) |
| n1_1 | $0.697^{***}$ (0.090) | $0.693^{***}$ (0.089) | $0.683^{***}$ (0.089) | $0.666^{***}$ (0.089) | $0.681^{***}$ (0.089) | $0.684^{***}$ (0.089) |
| n3_1 | $0.208^{***}$ (0.040) | $0.205^{***}$ (0.040) | $0.198^{***}$ (0.040) | $0.201^{***}$ (0.040) | $0.203^{***}$ (0.040) | $0.201^{***}$ (0.040) |
| n5_1 | $0.166^{***}$ (0.029) | $0.174^{***}$ (0.029) | $0.187^{***}$ (0.029) | $0.180^{***}$ (0.029) | $0.184^{***}$ (0.029) | $0.185^{***}$ (0.029) |
| n10_1 | $-0.080^{***}$ (0.020) | $-0.078^{***}$ (0.020) | $-0.082^{***}$ (0.019) | $-0.080^{***}$ (0.019) | $-0.079^{***}$ (0.019) | $-0.080^{***}$ (0.019) |
| n15_1 | $0.066^{***}$ (0.013) | $0.065^{***}$ (0.013) | $0.064^{***}$ (0.013) | $0.066^{***}$ (0.013) | $0.065^{***}$ (0.013) | $0.064^{***}$ (0.013) |
| nAll_2 | $-0.002^{**}$ (0.001) | $-0.002^{**}$ (0.001) | $-0.002^{**}$ (0.001) | $-0.002^{**}$ (0.001) | $-0.002^{**}$ (0.001) | $-0.002^{**}$ (0.001) |
| n1_2 | $-0.311^{***}$ (0.054) | $-0.316^{***}$ (0.054) | $-0.309^{***}$ (0.054) | $-0.308^{***}$ (0.054) | $-0.313^{***}$ (0.054) | $-0.312^{***}$ (0.054) |
| n10_2 | $0.066^{***}$ (0.010) | $0.065^{***}$ (0.010) | $0.065^{***}$ (0.010) | $0.066^{***}$ (0.010) | $0.066^{***}$ (0.010) | $0.066^{***}$ (0.010) |
| n15_2 | $-0.033^{***}$ (0.007) | $-0.031^{***}$ (0.007) | $-0.030^{***}$ (0.007) | $-0.032^{***}$ (0.007) | $-0.031^{***}$ (0.007) | $-0.030^{***}$ (0.007) |
| nAll_3 | $-0.018^{***}$ (0.001) | $-0.017^{***}$ (0.001) | $-0.018^{***}$ (0.001) | $-0.018^{***}$ (0.001) | $-0.017^{***}$ (0.001) | $-0.017^{***}$ (0.001) |
| n1_3 | $0.384^{***}$ (0.061) | $0.379^{***}$ (0.061) | $0.376^{***}$ (0.060) | $0.375^{***}$ (0.060) | $0.368^{***}$ (0.060) | $0.370^{***}$ (0.060) |
| n3_3 | $-0.011$ (0.034) | $-0.015$ (0.034) | $-0.013$ (0.034) | $-0.012$ (0.034) | $-0.013$ (0.033) | $-0.012$ (0.033) |
| n5_3 | $-0.120^{***}$ (0.021) | $-0.117^{***}$ (0.021) | $-0.119^{***}$ (0.021) | $-0.122^{***}$ (0.021) | $-0.119^{***}$ (0.021) | $-0.120^{***}$ (0.021) |
| n10_3 | $0.086^{***}$ (0.007) | $0.085^{***}$ (0.007) | $0.086^{***}$ (0.007) | $0.086^{***}$ (0.007) | $0.086^{***}$ (0.007) | $0.086^{***}$ (0.007) |
| nAll_4 | $-0.007^{***}$ (0.001) | $-0.006^{***}$ (0.001) | $-0.007^{***}$ (0.001) | $-0.006^{***}$ (0.001) | $-0.006^{***}$ (0.001) | $-0.006^{***}$ (0.001) |
| n3_4 | $0.151^{***}$ (0.019) | $0.148^{***}$ (0.019) | $0.144^{***}$ (0.019) | $0.147^{***}$ (0.019) | $0.146^{***}$ (0.019) | $0.145^{***}$ (0.019) |
| n15_4 | $-0.014^{***}$ (0.005) | $-0.016^{***}$ (0.005) | $-0.015^{***}$ (0.005) | $-0.015^{***}$ (0.005) | $-0.016^{***}$ (0.005) | $-0.016^{***}$ (0.005) |
| cInt1: BB intercept angle | $0.002^{***}$ (0.001) | $0.002^{***}$ (0.001) | $0.002^{***}$ (0.001) | $0.002^{***}$ (0.001) | $0.002^{***}$ (0.001) | $0.002^{***}$ (0.001) |
| cInt3: SB2 intercept angle | $0.001^{*}$ (0.0005) | $0.001^{*}$ (0.0005) | $0.001^{*}$ (0.0005) | $0.001^{**}$ (0.0005) | $0.001^{*}$ (0.0005) | $0.001^{*}$ (0.0005) |
| ccbar_1 (BB-BB comp. center km) | $-0.013^{***}$ (0.004) | $-0.014^{***}$ (0.004) | $-0.014^{***}$ (0.004) | $-0.014^{***}$ (0.004) | $-0.014^{***}$ (0.004) | $-0.014^{***}$ (0.004) |
| ccbar_2 (SB1-SB1 comp. center km) | $0.018^{**}$ (0.008) | $0.019^{**}$ (0.008) | $0.018^{**}$ (0.007) | $0.018^{**}$ (0.007) | $0.017^{**}$ (0.007) | $0.017^{**}$ (0.007) |
| icbar_2 (avg. customer-SB1 km) | $-0.002$ (0.010) | $-0.002$ (0.010) | $-0.0003$ (0.010) | $-0.001$ (0.010) | 0.001 (0.010) | 0.0005 (0.010) |
| icbar_3 (avg. customer-SB2 km) | $0.009^{***}$ (0.003) | $0.009^{***}$ (0.003) | $0.009^{***}$ (0.003) | $0.009^{***}$ (0.003) | $0.009^{***}$ (0.003) | $0.009^{***}$ (0.003) |
| icbar_4 (avg. customer-SS km) | $-0.007$ (0.006) | $-0.007$ (0.006) | $-0.007$ (0.006) | $-0.007$ (0.006) | $-0.007$ (0.006) | $-0.008$ (0.006) |
| Trips/month (pre LP) | $0.015^{***}$ (0.005) | $0.015^{***}$ (0.005) | $0.016^{***}$ (0.005) | $0.016^{***}$ (0.005) | $0.016^{***}$ (0.005) | $0.016^{***}$ (0.005) |
| Total sales (pre LP) | $0.0002^{***}$ (0.00001) | $0.0002^{***}$ (0.00001) | $0.0002^{***}$ (0.00001) | $0.0002^{***}$ (0.00001) | $0.0002^{***}$ (0.00001) | $0.0002^{***}$ (0.00001) |
| Number of focal stores w/i 60km | $0.008^{***}$ (0.002) | $0.007^{***}$ (0.002) | $0.007^{***}$ (0.002) | $0.007^{***}$ (0.002) | $0.007^{***}$ (0.002) | $0.007^{***}$ (0.002) |
| Population Density x ic3 | $0.00000^{**}$ (0.00000) | $0.00000^{**}$ (0.00000) | $0.00000^{**}$ (0.00000) | $0.00000^{**}$ (0.00000) | $0.00000^{**}$ (0.00000) | $0.00000^{**}$ (0.00000) |
| Population Density x sc1 | 0.00000 (0.00000) | 0.00000 (0.00000) | 0.00000 (0.00000) | 0.00000 (0.00000) | 0.00000 (0.00000) | 0.00000 (0.00000) |
| Distinct SKUs/Basket | $-0.015$ (0.026) | $-0.013$ (0.026) | $-0.001$ (0.025) | $-0.0004$ (0.025) | 0.001 (0.025) | 0.0001 (0.025) |
| Distinct Categories/Basket | $0.004^{**}$ (0.002) | $0.004^{**}$ (0.002) | 0.002 (0.002) | 0.002 (0.002) | 0.002 (0.002) | 0.002 (0.002) |
| Number of Items/Basket | $-0.113^{**}$ (0.056) | $-0.115^{**}$ (0.056) | $-0.077$ (0.054) | $-0.075$ (0.054) | $-0.076$ (0.053) | $-0.076$ (0.053) |
| Prop. sales in LP cat. (pre LP) | $5.948^{***}$ (0.274) | $5.941^{***}$ (0.273) | | | | |
| Received promotional email | $0.140^{***}$ (0.031) | | $0.116^{***}$ (0.030) | | | |
| Received promotional email for LP | $1.920^{***}$ (0.226) | | | $1.468^{***}$ (0.206) | | |
| Days to LP category purchase | $-0.044^{***}$ (0.008) | | | | $-0.032^{***}$ (0.006) | |
| Days to LP discount purchase | 0.003 (0.005) | | | | | $-0.009^{**}$ (0.004) |
| Observations | 163,629 | 163,629 | 163,629 | 163,629 | 163,629 | 163,629 |
| Log Likelihood | −25,872.350 | −25,936.830 | −26,235.590 | −26,224.000 | −26,219.830 | −26,238.750 |
| Akaike Inf. Crit. | 51,840.700 | 51,961.670 | 52,559.180 | 52,535.990 | 52,527.650 | 52,565.510 |
| *Note:* | | | | | | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 |

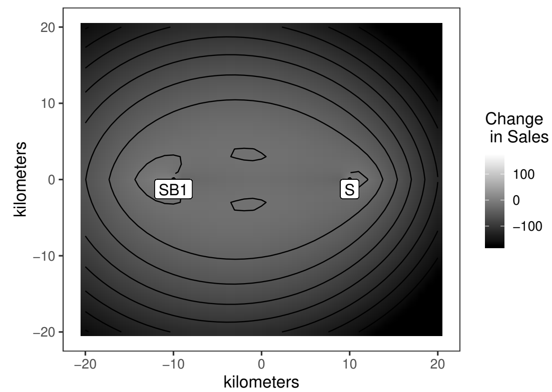
Table A.4: Post-LASSO Exclusion Variable Regressions

# Appendix 4: Additional Heatmaps

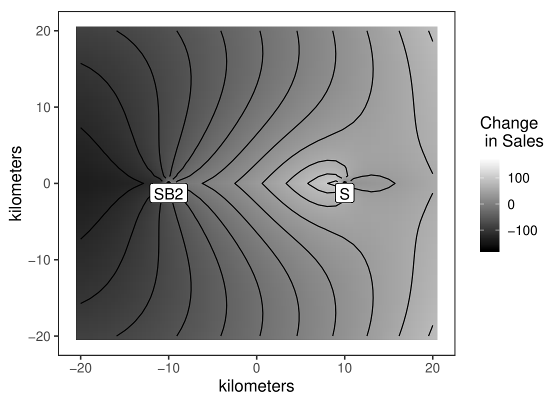
Figure A.2 shows additional heatmaps of interest.



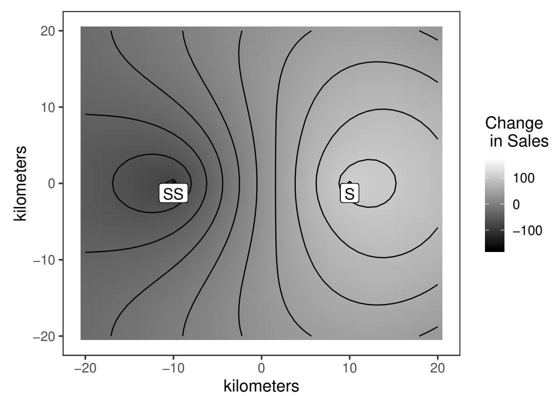

(a) Big Box (BB)




(b) Small Box #1 (SB1)




(c) Small Box #2 (SB2)




(d) Small Box Specialist (SS)

Figure A.2: Change in Monthly Spend by Competitive Structure